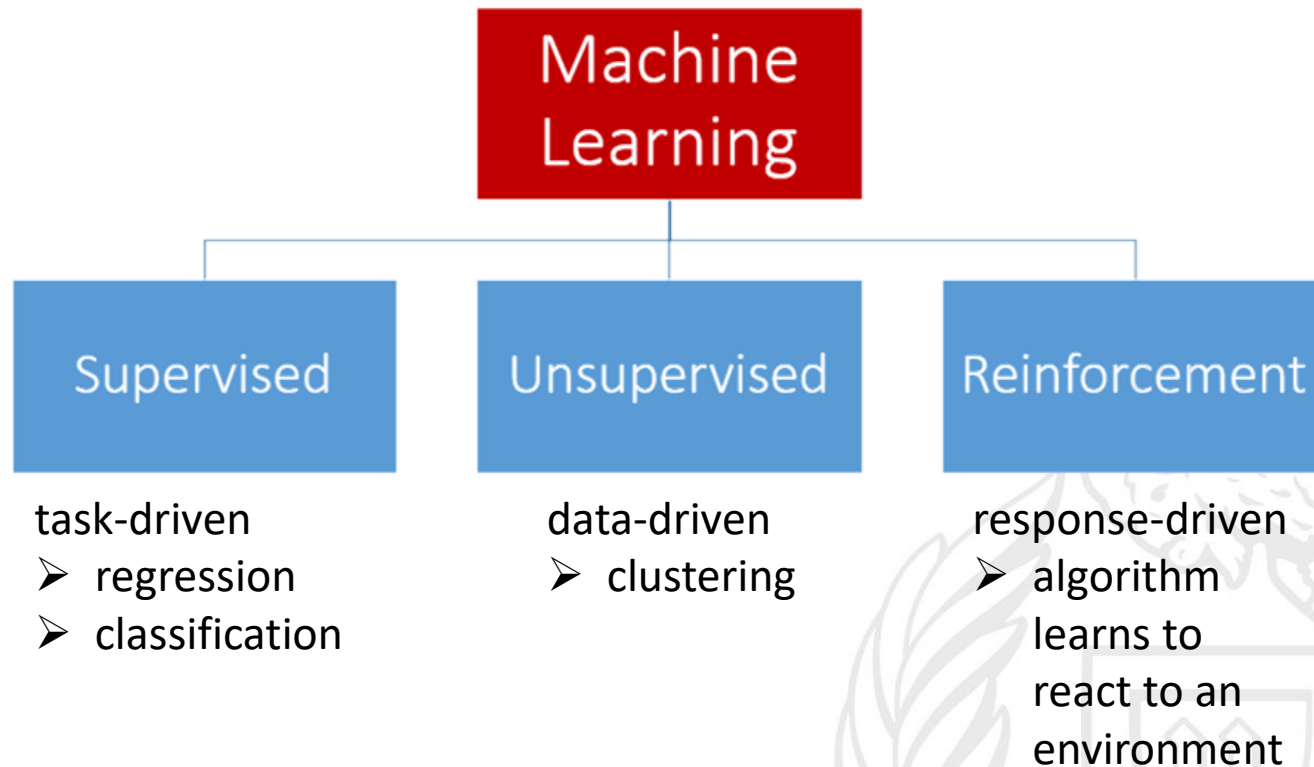
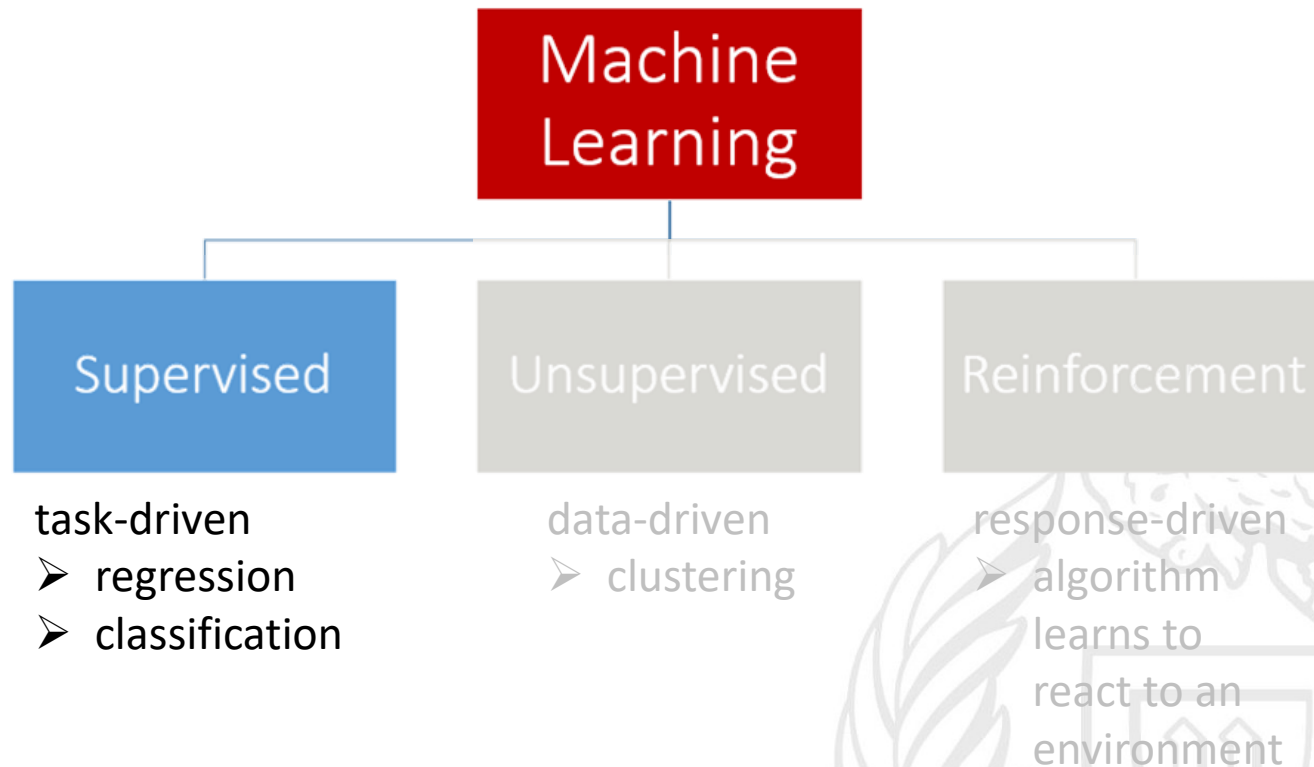


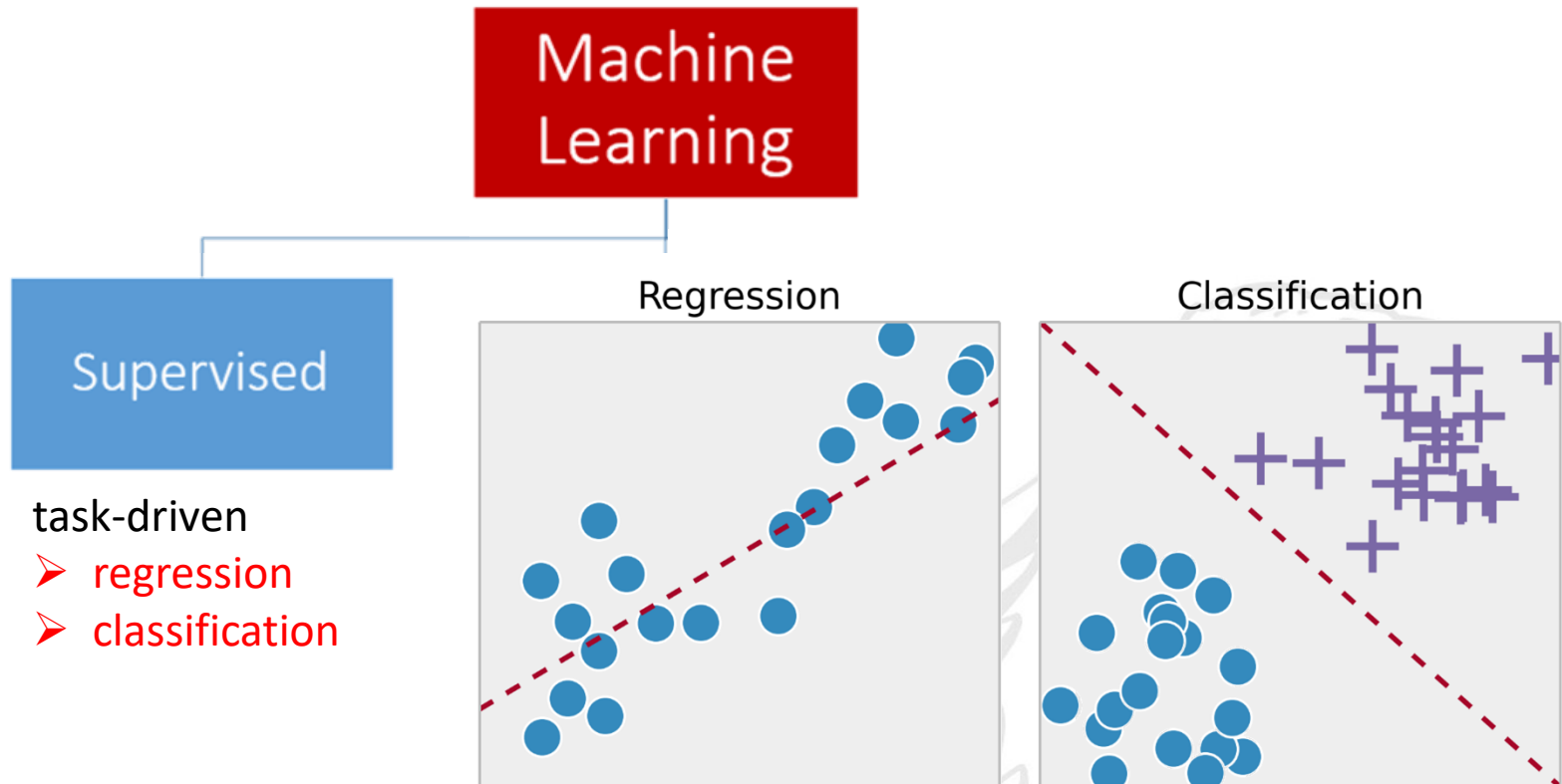
Types of machine learning:



Types of machine learning:



Types of machine learning:



ML as a sophisticated version of statistics

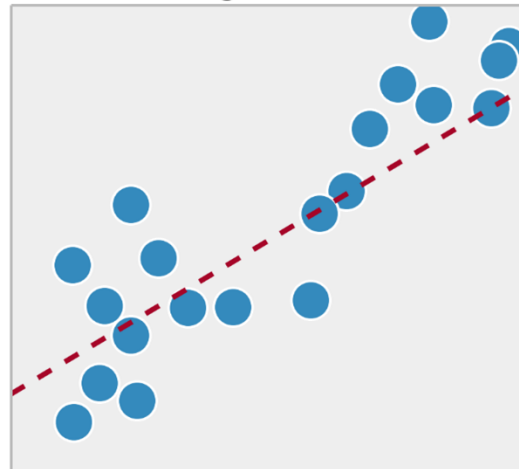


University at Buffalo

The State University of New York



Regression

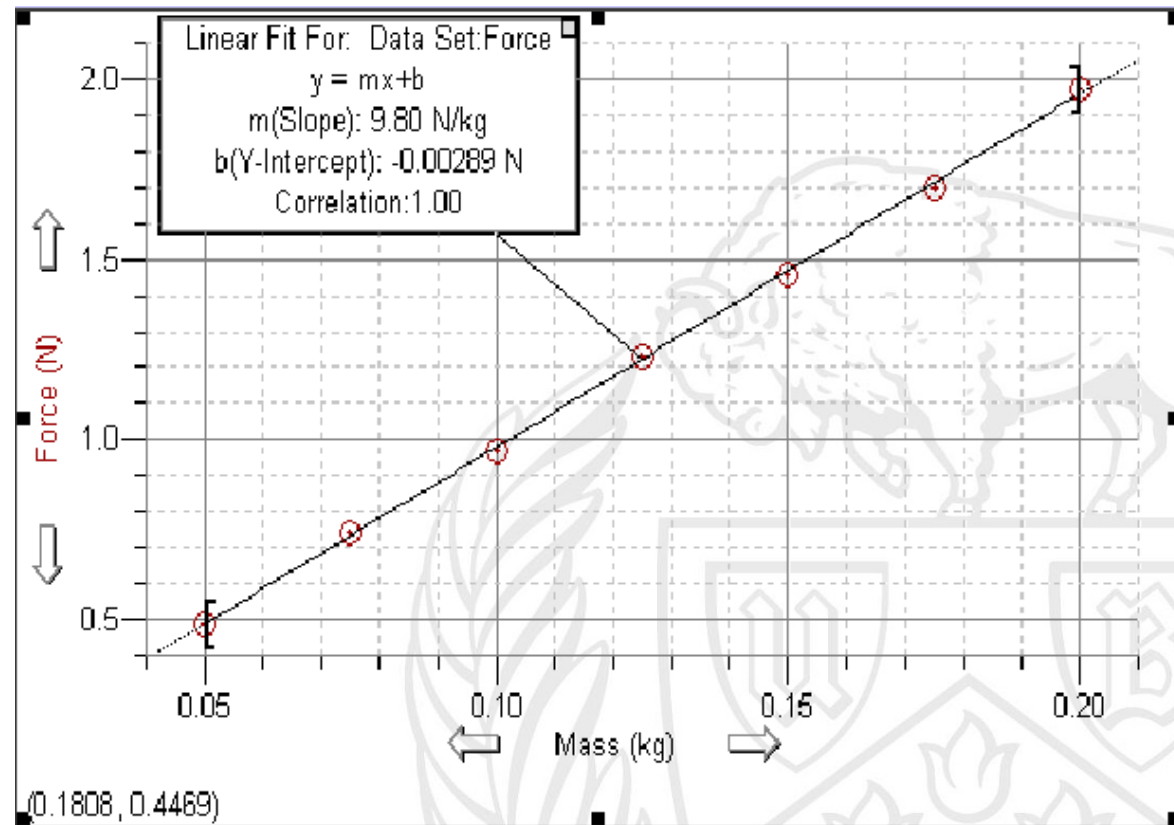


ML as a sophisticated version of statistics



University at Buffalo

The State University of New York





	force
n apples	...
	p_n
	...



	output
n input values	... p_n ...



	m features
n input values	
	 $d_{n,m}$
	
	

$$\begin{array}{c} \text{ML} \\ \longrightarrow \\ p_n = f(d_{n,m}) \end{array}$$

	output
n input values	...
	p_n
	...
	...



	m features ($m, V, \dots$)
n apples	
	 $d_{n,m}$
	
	

$$\xrightarrow{\text{ML}} \\ p_n = f(d_{n,m})$$

	force
n apples	...
	p_n
	...
	...



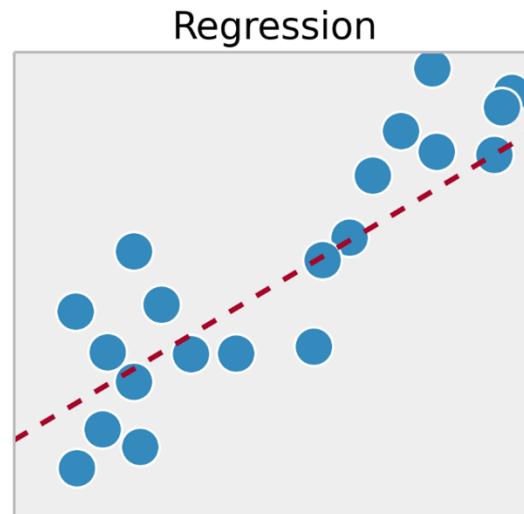
	m features
n input values	
	 $d_{n,m}$
	
	

$$\xrightarrow{\text{ML}} \\ p_n = f(d_{n,m})$$

	output
n input values	...
	p_n
	...
	...

“X”

“y”



Important machine learning techniques



University at Buffalo

The State University of New York

Method	Prediction Model	Optimization Problem
Linear Regression	$f(\mathbf{x}) = \mathbf{x}'\mathbf{w} + b$	$\min_{\mathbf{w}} = \frac{1}{n} \sum_{i=1}^n \ell_{\text{se}}(f(\mathbf{x}_i), y_i) \quad (\text{A.2})$
Ridge Regression	$f(\mathbf{x}) = \mathbf{x}'\mathbf{w} + b$	$\min_{\mathbf{w}} = \frac{1}{n} \sum_{i=1}^n \ell_{\text{se}}(f(\mathbf{x}_i), y_i) + \lambda \cdot \ \mathbf{w}\ ^2 \quad (\text{A.3})$
Kernel Ridge Regression	$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b$	$\min_{\alpha} = \frac{1}{n} \sum_{i=1}^n \ell_{\text{se}}(f(\mathbf{x}_i), y_i) + \lambda \cdot \sum_{i,j} \alpha_i k(\mathbf{x}_i, \mathbf{x}_j) \alpha_j \quad (\text{A.4})$
Support Vector Regression	$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b$	$\min_{\alpha} = \frac{1}{n} \sum_{i=1}^n \ell_{\varepsilon}(f(\mathbf{x}_i), y_i) + \lambda \cdot \sum_{i,j} \alpha_i k(\mathbf{x}_i, \mathbf{x}_j) \alpha_j \quad (\text{A.5})$
Support Vector Classification	$c(\mathbf{x}) = \text{sgn}(f(\mathbf{x}))$ with $f(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b$	$\min_{\alpha} = \frac{1}{n} \sum_{i=1}^n \ell_{+}(f(\mathbf{x}_i), y_i) + \lambda \cdot \sum_{i,j} \alpha_i k(\mathbf{x}_i, \mathbf{x}_j) \alpha_j \quad (\text{A.6})$

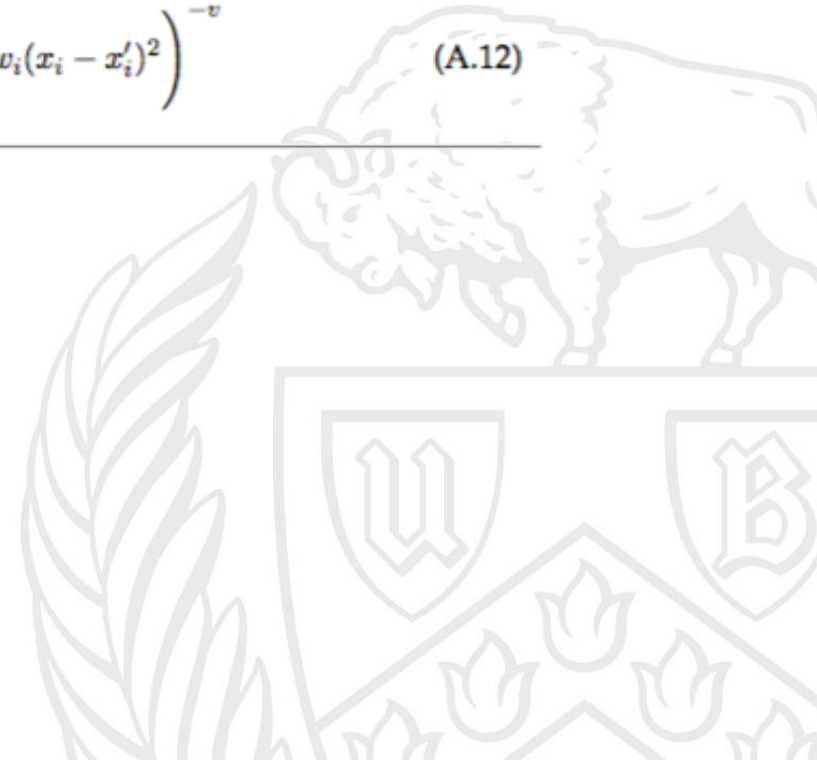
Important machine learning techniques



University at Buffalo

The State University of New York

Kernel Functions	
linear function	$k(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^d x_i x'_i$ (A.10)
radial basis function (RBF)	$k(\mathbf{x}, \mathbf{x}') = \exp \left(\frac{-\ \mathbf{x} - \mathbf{x}'\ ^2}{2\sigma^2} \right)$ (A.11)
rational quadratic function	$k(\mathbf{x}, \mathbf{x}') = \left(1 + \sum_{i=1}^d w_i (x_i - x'_i)^2 \right)^{-v}$ (A.12)



Important machine learning techniques: ANNs



University at Buffalo

The State University of New York

“Neural Networks Demystified” by Stephen Welch

<http://www.youtube.com/watch?v=bxe2T-V8XRs>

<http://github.com/stephencwelch/Neural-Networks-Demystified>

<http://www.welchlabs.com>



Neural Networks Demystified + Deep Learning Lectures

Neural Networks Demystified [Part 1: Data and Architecture]

```
In [1]: x = np.array([[3, 5], [5, 1], [10, 2]], dtype=float)
        y = np.array([75, 82, 93], dtype=float)

In [2]: x
Out[2]: array([[ 3.,  5.],
               [ 5.,  1.],
               [10.,  2.]])

In [3]: y
Out[3]: array([75.,
               82.,
               93.])

In [ ]:
```

Neural Networks Demystified [Part 2: Forward Propagation]

Neural Networks Demystified [Part 3: Gradient Descent]

Neural Networks Demystified [Part 4: Backpropagation]

Neural Networks Demystified [Part 5: Numerical Gradient Checking]

Neural Networks Demystified [Part 6: Training]

Neural Networks Demystified [Part 7: Overfitting, Testing, and Regularization]

Neural Networks Demystified [Part 8: Final Thoughts]

Neural Networks Demystified [Part 9: Conclusion]

Neural Networks Demystified [Part 10: Appendix]

Neural Networks Demystified [Part 11: License]

Neural Networks Demystified [Part 12: Acknowledgments]

Neural Networks Demystified [Part 13: Contact Information]

Neural Networks Demystified [Part 14: About the Author]

Neural Networks Demystified [Part 15: About the Series]

Neural Networks Demystified [Part 16: About the Book]

Neural Networks Demystified [Part 17: About the Publisher]

Neural Networks Demystified [Part 18: About the Series]

Neural Networks Demystified [Part 19: About the Book]

Neural Networks Demystified [Part 20: About the Publisher]

Neural Networks Demystified [Part 21: About the Series]

Neural Networks Demystified [Part 22: About the Book]

Neural Networks Demystified [Part 23: About the Publisher]

Neural Networks Demystified [Part 24: About the Series]

Neural Networks Demystified [Part 25: About the Book]

Neural Networks Demystified [Part 26: About the Publisher]

Neural Networks Demystified [Part 27: About the Series]

Neural Networks Demystified [Part 28: About the Book]

Neural Networks Demystified [Part 29: About the Publisher]

Neural Networks Demystified [Part 30: About the Series]

Neural Networks Demystified [Part 31: About the Book]

Neural Networks Demystified [Part 32: About the Publisher]

Neural Networks Demystified [Part 33: About the Series]

Neural Networks Demystified [Part 34: About the Book]

Neural Networks Demystified [Part 35: About the Publisher]

Neural Networks Demystified [Part 36: About the Series]

Neural Networks Demystified [Part 37: About the Book]

Neural Networks Demystified [Part 38: About the Publisher]

Neural Networks Demystified [Part 39: About the Series]

Neural Networks Demystified [Part 40: About the Book]

Neural Networks Demystified [Part 41: About the Publisher]

Neural Networks Demystified [Part 42: About the Series]

Neural Networks Demystified [Part 43: About the Book]

Neural Networks Demystified [Part 44: About the Publisher]

Neural Networks Demystified [Part 45: About the Series]

Neural Networks Demystified [Part 46: About the Book]

Neural Networks Demystified [Part 47: About the Publisher]

Neural Networks Demystified [Part 48: About the Series]

Neural Networks Demystified [Part 49: About the Book]

Neural Networks Demystified [Part 50: About the Publisher]

Neural Networks Demystified [Part 51: About the Series]

Neural Networks Demystified [Part 52: About the Book]

Neural Networks Demystified [Part 53: About the Publisher]

Neural Networks Demystified [Part 54: About the Series]

Neural Networks Demystified [Part 55: About the Book]

Neural Networks Demystified [Part 56: About the Publisher]

Neural Networks Demystified [Part 57: About the Series]

Neural Networks Demystified [Part 58: About the Book]

Neural Networks Demystified [Part 59: About the Publisher]

Neural Networks Demystified [Part 60: About the Series]

Neural Networks Demystified [Part 61: About the Book]

Neural Networks Demystified [Part 62: About the Publisher]

Neural Networks Demystified [Part 63: About the Series]

Neural Networks Demystified [Part 64: About the Book]

Neural Networks Demystified [Part 65: About the Publisher]

Neural Networks Demystified [Part 66: About the Series]

Neural Networks Demystified [Part 67: About the Book]

Neural Networks Demystified [Part 68: About the Publisher]

Neural Networks Demystified [Part 69: About the Series]

Neural Networks Demystified [Part 70: About the Book]

Neural Networks Demystified [Part 71: About the Publisher]

Neural Networks Demystified [Part 72: About the Series]

Neural Networks Demystified [Part 73: About the Book]

Neural Networks Demystified [Part 74: About the Publisher]

Neural Networks Demystified [Part 75: About the Series]

Neural Networks Demystified [Part 76: About the Book]

Neural Networks Demystified [Part 77: About the Publisher]

Neural Networks Demystified [Part 78: About the Series]

Neural Networks Demystified [Part 79: About the Book]

Neural Networks Demystified [Part 80: About the Publisher]

Neural Networks Demystified [Part 81: About the Series]

Neural Networks Demystified [Part 82: About the Book]

Neural Networks Demystified [Part 83: About the Publisher]

Neural Networks Demystified [Part 84: About the Series]

Neural Networks Demystified [Part 85: About the Book]

Neural Networks Demystified [Part 86: About the Publisher]

Neural Networks Demystified [Part 87: About the Series]

Neural Networks Demystified [Part 88: About the Book]

Neural Networks Demystified [Part 89: About the Publisher]

Neural Networks Demystified [Part 90: About the Series]

Neural Networks Demystified [Part 91: About the Book]

Neural Networks Demystified [Part 92: About the Publisher]

Neural Networks Demystified [Part 93: About the Series]

Neural Networks Demystified [Part 94: About the Book]

Neural Networks Demystified [Part 95: About the Publisher]

Neural Networks Demystified [Part 96: About the Series]

Neural Networks Demystified [Part 97: About the Book]

Neural Networks Demystified [Part 98: About the Publisher]

Neural Networks Demystified [Part 99: About the Series]

Neural Networks Demystified [Part 100: About the Book]

Neural Networks Demystified [Part 101: About the Publisher]

Neural Networks Demystified [Part 102: About the Series]

Neural Networks Demystified [Part 103: About the Book]

Neural Networks Demystified [Part 104: About the Publisher]

Neural Networks Demystified [Part 105: About the Series]

Neural Networks Demystified [Part 106: About the Book]

Neural Networks Demystified [Part 107: About the Publisher]

Neural Networks Demystified [Part 108: About the Series]

Neural Networks Demystified [Part 109: About the Book]

Neural Networks Demystified [Part 110: About the Publisher]

Neural Networks Demystified [Part 111: About the Series]

Neural Networks Demystified [Part 112: About the Book]

Neural Networks Demystified [Part 113: About the Publisher]

Neural Networks Demystified [Part 114: About the Series]

Neural Networks Demystified [Part 115: About the Book]

Neural Networks Demystified [Part 116: About the Publisher]

Neural Networks Demystified [Part 117: About the Series]

Neural Networks Demystified [Part 118: About the Book]

Neural Networks Demystified [Part 119: About the Publisher]

Neural Networks Demystified [Part 120: About the Series]

Neural Networks Demystified [Part 121: About the Book]

Neural Networks Demystified [Part 122: About the Publisher]

Neural Networks Demystified [Part 123: About the Series]

Neural Networks Demystified [Part 124: About the Book]

Neural Networks Demystified [Part 125: About the Publisher]

Neural Networks Demystified [Part 126: About the Series]

Neural Networks Demystified [Part 127: About the Book]

Neural Networks Demystified [Part 128: About the Publisher]

Neural Networks Demystified [Part 129: About the Series]

Neural Networks Demystified [Part 130: About the Book]

Neural Networks Demystified [Part 131: About the Publisher]

Neural Networks Demystified [Part 132: About the Series]

Neural Networks Demystified [Part 133: About the Book]

Neural Networks Demystified [Part 134: About the Publisher]

Neural Networks Demystified [Part 135: About the Series]

Neural Networks Demystified [Part 136: About the Book]

Neural Networks Demystified [Part 137: About the Publisher]

Neural Networks Demystified [Part 138: About the Series]

Neural Networks Demystified [Part 139: About the Book]

Neural Networks Demystified [Part 140: About the Publisher]

Neural Networks Demystified [Part 141: About the Series]

Neural Networks Demystified [Part 142: About the Book]

Neural Networks Demystified [Part 143: About the Publisher]

Neural Networks Demystified [Part 144: About the Series]

Neural Networks Demystified [Part 145: About the Book]

Neural Networks Demystified [Part 146: About the Publisher]

Neural Networks Demystified [Part 147: About the Series]

Neural Networks Demystified [Part 148: About the Book]

Neural Networks Demystified [Part 149: About the Publisher]

Neural Networks Demystified [Part 150: About the Series]

Neural Networks Demystified [Part 151: About the Book]

Neural Networks Demystified [Part 152: About the Publisher]

Neural Networks Demystified [Part 153: About the Series]

Neural Networks Demystified [Part 154: About the Book]

Neural Networks Demystified [Part 155: About the Publisher]

Neural Networks Demystified [Part 156: About the Series]

Neural Networks Demystified [Part 157: About the Book]

Neural Networks Demystified [Part 158: About the Publisher]

Neural Networks Demystified [Part 159: About the Series]

Neural Networks Demystified [Part 160: About the Book]

Neural Networks Demystified [Part 161: About the Publisher]

Neural Networks Demystified [Part 162: About the Series]

Neural Networks Demystified [Part 163: About the Book]

Neural Networks Demystified [Part 164: About the Publisher]

Neural Networks Demystified [Part 165: About the Series]

Neural Networks Demystified [Part 166: About the Book]

Neural Networks Demystified [Part 167: About the Publisher]

Neural Networks Demystified [Part 168: About the Series]

Neural Networks Demystified [Part 169: About the Book]

Neural Networks Demystified [Part 170: About the Publisher]

Neural Networks Demystified [Part 171: About the Series]

Neural Networks Demystified [Part 172: About the Book]

Neural Networks Demystified [Part 173: About the Publisher]

Neural Networks Demystified [Part 174: About the Series]

Neural Networks Demystified [Part 175: About the Book]

Neural Networks Demystified [Part 176: About the Publisher]

Neural Networks Demystified [Part 177: About the Series]

Neural Networks Demystified [Part 178: About the Book]

Neural Networks Demystified [Part 179: About the Publisher]

Neural Networks Demystified [Part 180: About the Series]

Neural Networks Demystified [Part 181: About the Book]

Neural Networks Demystified [Part 182: About the Publisher]

Neural Networks Demystified [Part 183: About the Series]

Neural Networks Demystified [Part 184: About the Book]

Neural Networks Demystified [Part 185: About the Publisher]

Neural Networks Demystified [Part 186: About the Series]

Neural Networks Demystified [Part 187: About the Book]

Neural Networks Demystified [Part 188: About the Publisher]

Neural Networks Demystified [Part 189: About the Series]

Neural Networks Demystified [Part 190: About the Book]

Neural Networks Demystified [Part 191: About the Publisher]

Neural Networks Demystified [Part 192: About the Series]

Neural Networks Demystified [Part 193: About the Book]

Neural Networks Demystified [Part 194: About the Publisher]

Neural Networks Demystified [Part 195: About the Series]

Neural Networks Demystified [Part 196: About the Book]

Neural Networks Demystified [Part 197: About the Publisher]

Neural Networks Demystified [Part 198: About the Series]

Neural Networks Demystified [Part 199: About the Book]

Neural Networks Demystified [Part 200: About the Publisher]

Personal Open source Business Explore Pricing Blog Support This repository Search Sign in Sign up

stephencwelch / Neural-Networks-Demystified

Watch 17 Star 270 Fork 92

Code Issues 1 Pull requests 2 Pulse Graphs

Supporting code for short YouTube series Neural Networks Demystified.

26 commits 1 branch 0 releases 2 contributors

Branch master New pull request New file Find file HTTPS https://github.com/stephencwelch/Neural-Networks-Demystified Download ZIP

stephencwelch Merge pull request #9 from ForceBru/patch-1 Latest commit 1d352e1 on Nov 27, 2015

File	Commit	Time
ipynb_checkpoints	Added %Spylab inline to beginning of notebooks	4 months ago
images	updated readme	a year ago
DS_Store	Added %Spylab inline to beginning of notebooks	4 months ago
Part 1 Data and Architecture.ipynb	Rename Part 1 Data and Architecture.ipynb to Part 1 Data and Archite...	4 months ago
Part 2 Forward Propagation.ipynb	Added %Spylab inline to beginning of notebooks	4 months ago
Part 3 Gradient Descent.ipynb	Added %Spylab inline to beginning of notebooks	4 months ago
Part 4 Backpropagation.ipynb	Added %Spylab inline to beginning of notebooks	4 months ago
Part 5 Numerical Gradient Checking.ipynb	Added %Spylab inline to beginning of notebooks	4 months ago
Part 6 Training.ipynb	Added %Spylab inline to beginning of notebooks	4 months ago
Part 7 Overfitting, Testing, and Regularization.ipynb	Added %Spylab inline to beginning of notebooks	4 months ago
README.md	Update README.md	4 months ago
Scratch.ipynb	finished part 7	a year ago
partFive.py	Rename PartFive.py to partFive.py	a year ago

Important machine learning techniques: ANNs

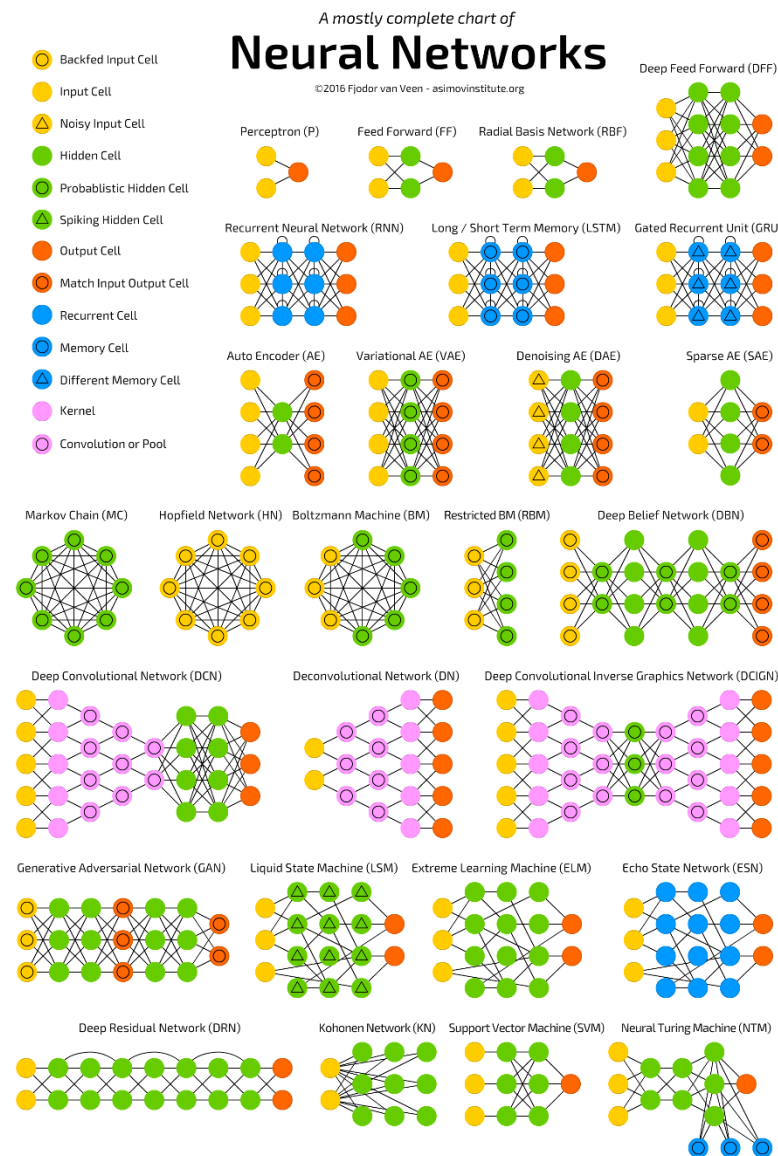
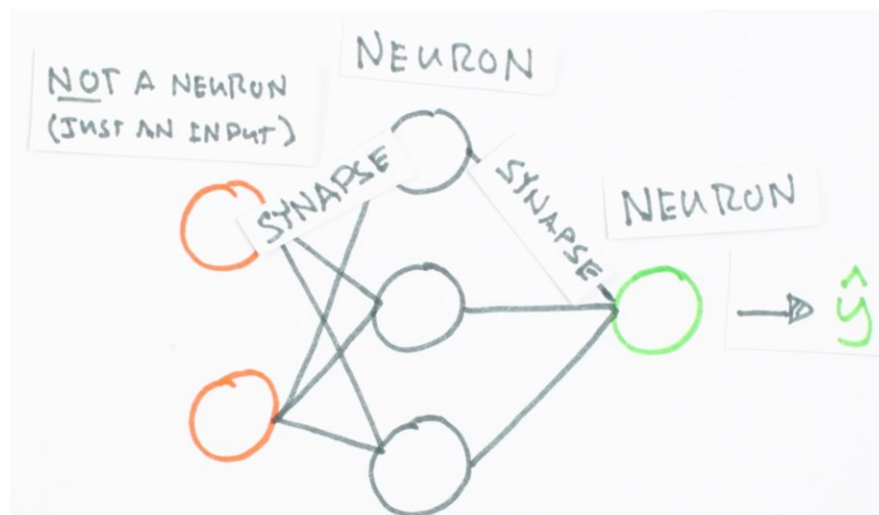


University at Buffalo

The State University of New York

“Neural Networks Demystified” by Stephen We

- ANNs from an algorithmic perspective
 - synapses (connect neurons, apply weight)
 - neurons (operation: sum synapses input)
 - note: complicated NN architectures possible



“Neural Networks Demystified” by Stephen Welch – take-home messages:

- ANNs from an algorithmic perspective
 - synapses (connect neurons, apply weights W)
 - neurons (operation: sum synapses inputs, apply activation)
 - note: complicated NN architectures possible (DNN, CNN, GAN, VAE, ...)
- ANNs as a linear algebra problem
- ANN training as an optimization problem (minimizing cost/error function wrt W)

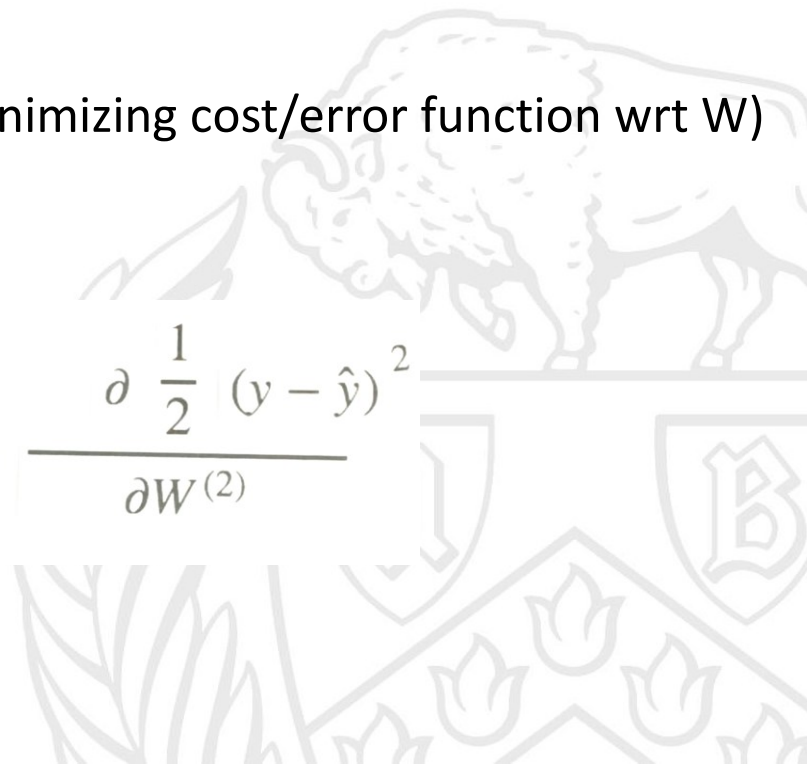
$$z^{(2)} = XW^{(1)} \quad (1)$$

$$a^{(2)} = f(z^{(2)}) \quad (2)$$

$$z^{(3)} = a^{(2)}W^{(2)} \quad (3)$$

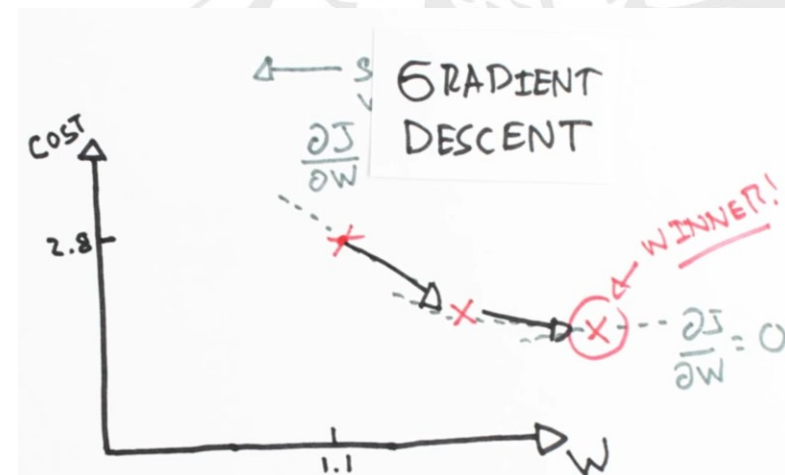
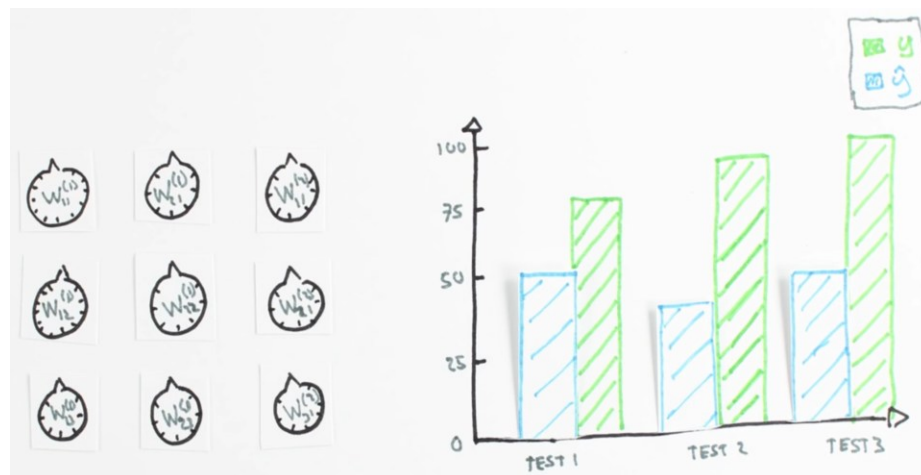
$$\hat{y} = f(z^{(3)}) \quad (4)$$

$$J = \sum \frac{1}{2} (y - \hat{y})^2 \quad (5)$$

$$\frac{\partial J}{\partial W^{(2)}} = \sum \frac{\partial \frac{1}{2} (y - \hat{y})^2}{\partial W^{(2)}}$$


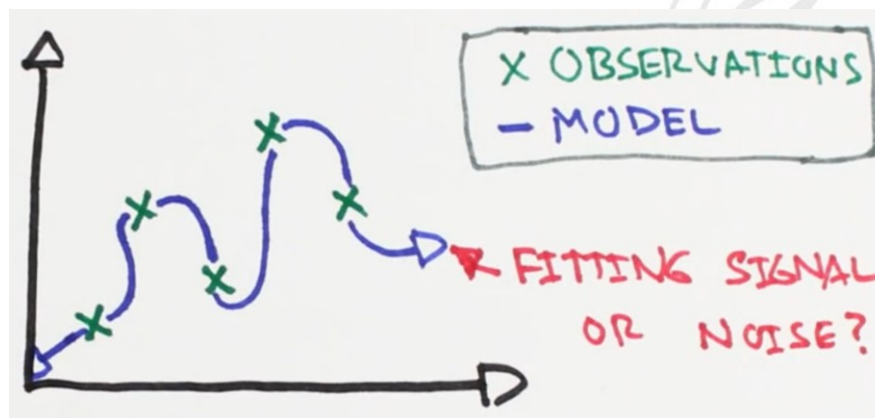
“Neural Networks Demystified” by Stephen Welch – take-home messages:

- ANNs from an algorithmic perspective
 - synapses (connect neurons, apply weights W)
 - neurons (operation: sum synapses inputs, apply activation)
 - note: complicated NN architectures possible (DNN, CNN, GAN, VAE, ...)
- ANNs as a linear algebra problem
- ANN training as an optimization problem (minimizing cost/error function wrt W)
- curse of dimensionality in brute-force methods: use gradient descent instead

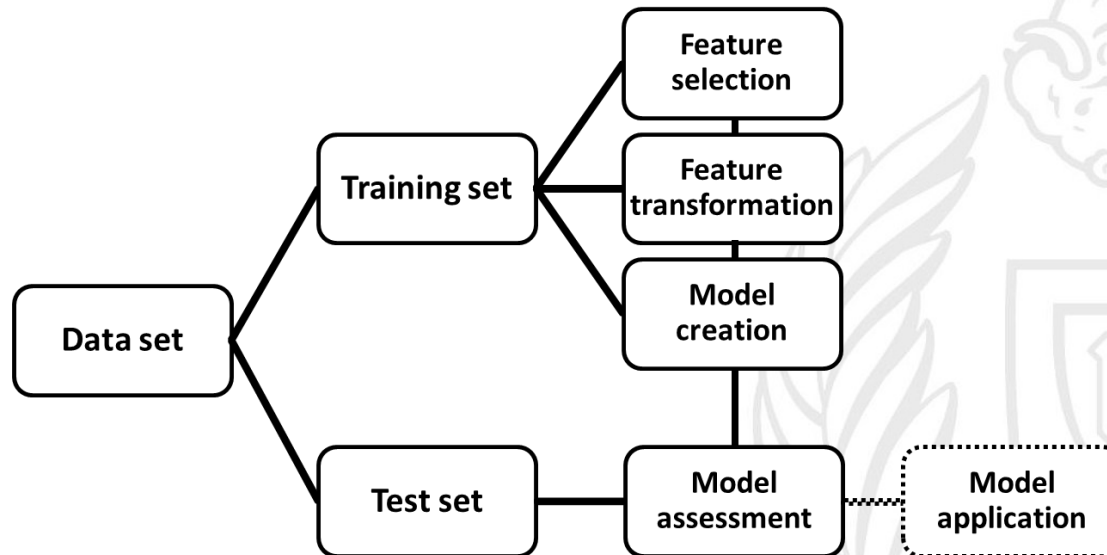
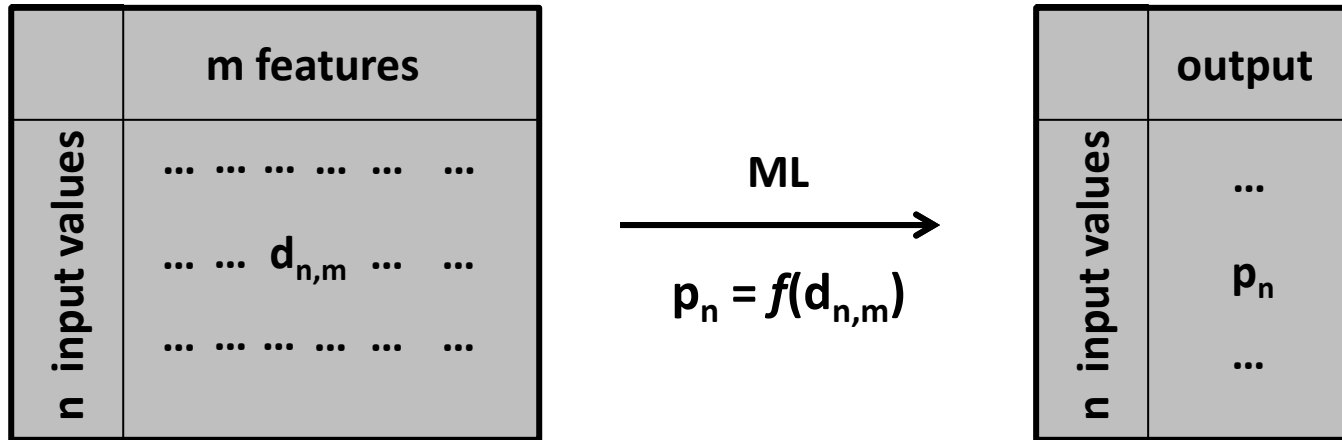


“Neural Networks Demystified” by Stephen Welch – take-home messages:

- ANNs from an algorithmic perspective
 - synapses (connect neurons, apply weights W)
 - neurons (operation: sum synapses inputs, apply activation)
 - note: complicated NN architectures possible (DNN, CNN, GAN, VAE, ...)
- ANNs as a linear algebra problem
- ANN training as an optimization problem (minimizing cost/error function wrt W)
- curse of dimensionality in brute-force methods: use gradient descent instead
- fitting the signal, not the noise: regularization to avoid overfitting

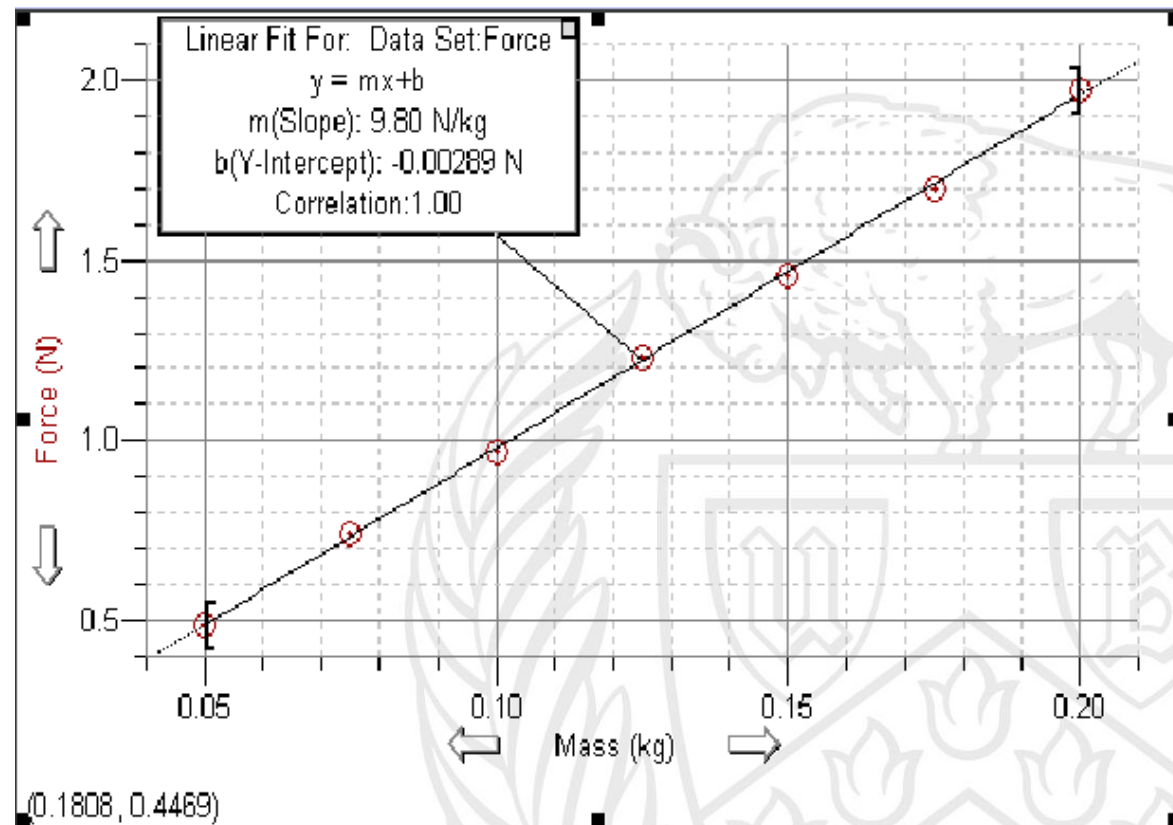


Important issues in the ML process



Important issues:

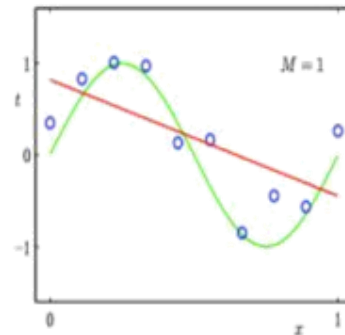
- suitable choice of features



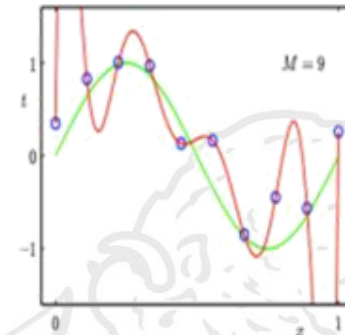
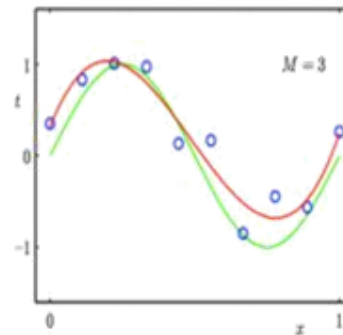
Important issues:

- suitable choice of features
- danger of under- or overfitting

Regression:

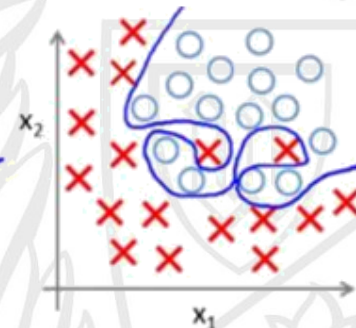
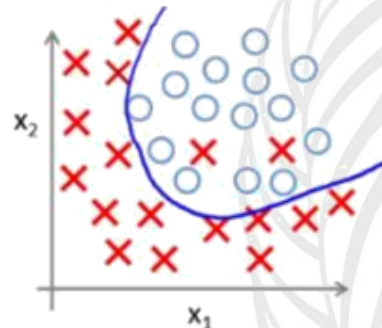
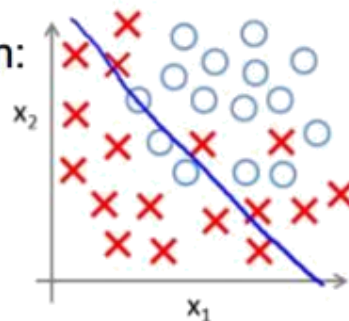


predictor too inflexible:
cannot capture pattern



predictor too flexible:
fits noise in the data

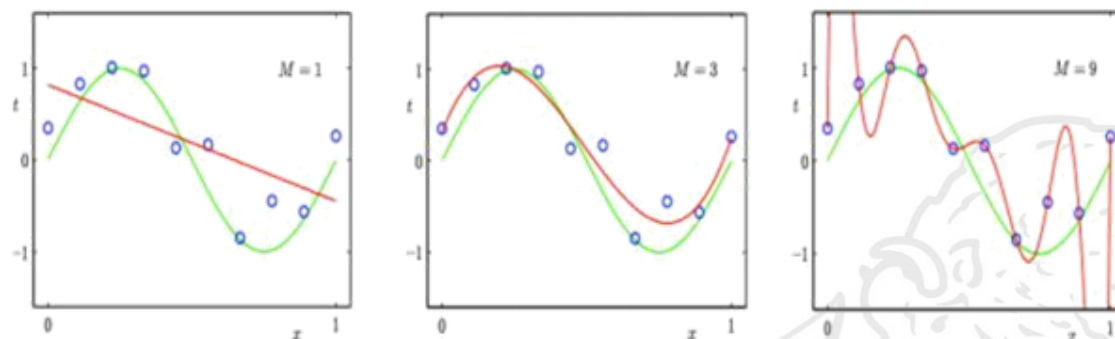
Classification:



Important issues:

- suitable choice of features
- danger of under- or overfitting

Regression:



Regularization Functions

squared error loss	$\ell_{se}(f(\mathbf{x}_i), y_i) = (y_i - f(\mathbf{x}_i))^2$ (A.7)
ε -intensive loss	$\ell_{\varepsilon}(f(\mathbf{x}_i), y_i) = f(\mathbf{x}_i) - y_i _{\varepsilon} = \begin{cases} 0 & \text{if } f(\mathbf{x}_i) - y_i \leq \varepsilon \\ f(\mathbf{x}_i) - y_i - \varepsilon & \text{else} \end{cases}$ (A.8)
hinge loss	$\ell_{+}(f(\mathbf{x}_i), y_i) = 1 - y_i f(\mathbf{x}_i) _{+} = \begin{cases} 0 & \text{if } 1 - y_i f(\mathbf{x}_i) \leq 0 \\ 1 - y_i f(\mathbf{x}_i) & \text{else} \end{cases}$ (A.9)

Important issues:

- suitable choice of features
- danger of under- or overfitting
- cross-validation against trainings bias

