A Comparative Study of Neural Networks and Logistic Regression for High Energy Physics

> Nigel Pugh, Joel Gonzalez-Santiago, and Thomas Johnson Mentor: Jerome E. Mitchell

Abstract

-- Collisions particle colliders are a source of particle discoveries. However, determining these particles requires solving difficult signal-versus-background classification problems

-- Other approaches have relied on machine learning models, which are limited in learning complex functions

-- Recent advances (computation and data size) neural networks provide an opportunity to learn complex functions and better discriminate between signal and background classes

-- We study the design considerations for a neural network model and provide a comparative analysis with logistic regression

Machine Learning Problems



The Machine Learning Framework



-- Training: given a *training set* of labeled examples $\{(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_N, \mathbf{y}_N)\}$, estimate the prediction function f by minimizing the prediction error on the training set

-- Testing: apply f to a never before seen *test example* \mathbf{x} and output the predicted value $y = f(\mathbf{x})$

Steps



Generalization



Training set (labels known)



Test set (labels unknown)

How well does a learned model generalize from the data it was trained on to a new test set?

Our Data...

Classification problem to distinguish between a signal process to background for high energy physics



Data Set Characteristics:	N/A	Number of Instances:	5000000	Area:	Physical
Attribute Characteristics:	Real	Number of Attributes:	18	Date Donated	2014-02-12
Associated Tasks:	Classification	Missing Values?	N/A	Number of Web Hits:	31151

	H	iome In	sent R	age Layout	Formu	las Dat	a Revi	ow. Viev	Add-	ins
	1	, X	Calibri	(Body) -	12 - /	A- A-	= =	- %	123	Wrap Text
	Pa	ete .	в	r u -		. <u>A</u> .	= =	-	1	Merge & O
			Y SZO	6 0.9728	614687919	61				
	HA.	*		F. 5.5720	D		1	6	н	
	* ⊼	0.785.01	6 54E-01	1.185400	1.1654000	1.748.400	-8 24F-01	5 68F /01	-1 75E-01	0.005+00
	51	1675+00	6 438 02	1 236400	5 268,07	3 395.07	1.676400	3.485+00	-1.225+00	1.005+00
		A 45E-01	-1.345-01	-7.105-01	4 528-01	1.615+00	-7 895-01	1.225+00	5.045-01	1.000+00
		3 818-01	-0 76E-01	6 935-01	4.495-01	8 97E-01	-6.775-01	2.035+00	1.538+00	1.005+00
		1,315+00	-6.905-01	-6.768-01	1.595+00	-6.936-01	6.236-01	1.095+00	-1.828-01	1.005+00
-eatures	6	4.56E-01	1.10E+00	1.516+00	7.528-01	6.395-01	-7.428-01	3.23E-01	1.325+00	0.005+00
cutures	7	4 38F-01	-1.12E+00	-1.34E+00	5.02E-01	-1.72E+00	1.02E+00	2.15E-01	-4.615-01	0.005+00
		5 778-01	2.06E-01	4.26E-01	5 79E-01	9.27E-01	-1.10E+00	1.17E-01	1.55E+00	0.005+00
	-14	2.11E+00	7.43E-01	-3.31E-01	8.05E-01	-2.89E-02	-1.45E+00	2.30E+00	1.45E+00	0 1.00E+00
	10	1.00E+00	-4.72E-01	5.56E-01	1,235+00	1.26E+00	-1.05E+00	4.38E-01	-1.33E+00	0.006+00
	5.5	8.40E-01	9.58E-01	-5.86E-01	1.32E+00	-1.06E+00	1.25E+00	2.62E-01	-2.73E-02	0.00E+00
	12	126E+00	7.336-01	3.31E-01	1.51E+00	3.71E-01	-1.59E+00	5.73E-01	-1.05E+00	0.00€+00
	19	6.988-01	1.69E+00	-1.13E+00	9.67E-01	1.50€+00	8.81E-01	2.43E-01	-2.296-01	0.00€+00
	14	5.78E-01	-6.90E-01	-3.90E-01	4.80E-01	-6.32E-01	1.21E+00	6.40E-01	-1.62E+00	0.00E+00
	15	7.98E-01	9.946-02	-1.10E+00	5.31E-01	-9.62E-01	1.50E+00	1.27E+00	6.758-01	1.00E+00
	26	4.57E-01	-4.85E-03	-1.37E+00	4.568-01	2.728-01	1.64E+00	8.36E-01	-2.556-01	1.000+00
	17	8.656-01	-1.115+00	1.385+00	1.236+00	-1.60E+00	-3.368-01	4.53E-01	1.408+00	0.000+000
tancos	2.8	6.798-01	-1.168-01	-1.458+00	7.268-01	8.335-01	-1.071+00	7.345-01	6.176-01	0.006+00
lances	19	6.97E-01	1.53E+00	4.521-01	9.545-01	1.09E+00	-1.26E+00	2,201-02	7.695-01	0.000+000
	20	1.01€+00	-2.196-01	1.27E+00	1.628+00	7.605-02	5.86E-01	5.618-01	1.71E-01	1.000+00
	21	8.38E-01	1.13E+00	-9.11E-01	6.17E-01	1.00E+00	1.29E+00	1.27E+00	-6.77E-01	1.00E+00
	22	1.64E+00	-1.83E+00	1.50E-01	2.70E+00	-1.30E-01	-1.38E+00	1.94E-01	-1.48E+00	0.00E+00
	23	1.47E+00	1.16E-01	-1.04E+00	8.77E-01	6.91E-01	113E+00	8.91E-01	4.27E-01	0.00€+00
	24	1.64E+00	-1.04E+00	5.43E-01	7.45E-01	3.08E-01	9.57E-01	1.25E+00	-1.48E+00	1.00E+00
	25	1.36E+00	-5.06E-01	-6.45E-01	1.05E+00	1.14E+00	6.99E-01	1.66E+00	-6.15E-02	1.00€+00
	26	1.30E+00	-2.55E-01	1.12E+00	9,798-01	-1.35E-01	-6.97E-01	1.03E+00	-1.66E-01	0.00E+00
CONTRACTOR DATE: N	27	3.796-01	+1.08E+00	1.01E+00	6.11E-01	-1.29E+00	-6.56E-02	9.16E-01	-8.86E-01	0.00E+00
Physical	28	6.76E-01	1.27E-01	9.56E-01	6E-01 6.86E-01 -1.99E-01 -1.71E+00 9.24E-01 -6.93E-01 0.00E+00					
2	29	4.64E-01	1.13E-02	3.69E-01	5.53E-01	3.89E-01	3.02E-01	1.66E+00	1.52E+00	0.00E+00
	30	1.16E+00	-1.22E+00	1.39E+00	1.29E+00	8.54E-01	-8.64E-01	5.52E-01	3.72E-01	0.000+00
014 02 12	31 8.00E-01 -9.84E-02 -7.95E-01 4.78E-01 -4.3	-4.30E-01	1.08E+00	8.88E-01	-3.40E-01	0.006+00				
014-02-12	32	3.41E+00	5.906-01	1.518+00	2.205+00	9.158-01	-2.448-01	1.00E+00	9.488-02	1.000+00
	33	1.43E+00	-2.73E-01	-2.77E-01	8.428-01	5.548-01	1.528+00	4.125+00	-4.128-01	1.005+00
4454	34	8.666-01	-1.32E+00	-1.47E+00	1.06E+00	-1.35E+00	5.82E-01	7.50E-0Z	1.628+00	0.000 +00
51151	14.00	2 2 2 2 2 2 2 2 2	7 1 6 6 04		10 10 10 10 10 10	C 2011 111				

Instances

Labels

Neural Networks

-- Neural networks are loosely based on the human brain

-- Utilized for classification purposes

-- Popularity of neural networks have increased



Backpropagation

- -- Common Method used to train Neural Network
- -- Supervised Learning Technique
- -- Reduces error respective to weights
- -- Activation function is utilized

Activation function	Equation	Example	1D Graph	
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant		
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant		
Linear	$\phi(z) = z$	Adaline, linear regression		
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \ge \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \le -\frac{1}{2}, \end{cases}$	Support vector machine		
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN		
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer NN	<i>.</i>	

Neural Networks



Error = $\frac{1}{2}$ (Target – Output_{OUT})²

Backpropagation



Neural Network Design Considerations

- -- What transfer function should be used?
- -- How many inputs does the network need?
- -- How many hidden neurons per hidden layer?
- -- How many outputs should the network have?

There is no standard methodology to determinate these values. Even there is some heuristic points, final values are determinate by a trial and error procedure

Logistic Regression

Assumes the following form for P(Y|X):

$$P(Y = 1|X) = \frac{1}{1 + \exp(-(w_0 + \sum_i w_i X_i))}$$

Logistic function applied to a linear function of the data

Logistic function (or Sigmoid):

$$+exp(-z)$$

or Sigmoid):





Logistic Regression Application

Vegetation along transect with reflectance



Reflectance along Transect

Experiment : Our Neural Network Information

- --1000 training examples
- -- 100 test examples
- -- 8 features
- -- 3 Layer Neural Network
- -- 10,000 training iterations
- -- Hidden Nodes: Range from 1 100
- -- Step size : .01

Results: Neural Networks



Results: Best hidden node NN vs Logistic Regression



Conclusion

-- Neural Network with 65 hidden nodes provided better performance than our Logistic Regression model on all test cases

-- We conclude it is better to use a the Backpropagation algorithm to classify the High Energy Physics Data

Future Work

-- Increase the number of hidden layers, to determine if accuracy percentage will increase

-- Compare against different machine learning algorithms against this same data, and determine which are more effective for classification.

References

Baldi, P., P. Sadowski, and D. Whiteson. "Searching For Exotic Particles In High-Energy Physics With Deep Learning". *Nature Communications* 5.4308 (2014): 9. Web. 13 Mar. 2017.

Sadowski, Peter, Julian Collando, and Pierre Baldi. "Deep Learning, Dark Knowledge, And Dark Matter". JMLR Workshop And Conference Proceedings 42. Irvine, California: Journal of Machine Learning Research, 2017. 81-97. Web. 12 Mar. 2017.

O. Ohrimenko, F. Schuster, C. Fournet, A. Mehta, S. Nowozin, K. Vaswani and M. Costa, "Oblivious Multi-Party Machine Learning on Trusted Processors", in 25th USENIX Security Symposium, Austin, Texas, 2016, pp. 618-636.

D. Marrón, J. Read, A. Bifet and N. Navarro, "Data stream classification using random feature functions and novel method combinations", *Journal of Systems and Software*, vol. 127, pp. 195-204, 2017.

D. Whiteson and M. Lichman, "UCI Machine Learning Repository: SUSY Data Set", *Archive.ics.uci.edu*, 2017. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/SUSY. [Accessed: 03- Apr- 2017].

Acknowledgements

The team would like to acknowledge:

-- Jerome Mitchell, for his guidance, and contributions in completing this research project

-- Dr. Linda Hayden, who provided funding and opportunity of this project through the CERSER program

Questions?