

NeuraMorph

Validation results

P. Baillehache

October 3, 2020

Contents

1	Breast Cancer Wisconsin (Diagnostic) Data Set	2
1.1	NeuraMorph results	2
1.2	Concurrent results	2
1.2.1	Multisurface method	2
1.2.2	Additive logistic regression	2
1.3	Validation result	3
2	Iris Data Set	4
2.1	NeuraMorph results	4
2.2	Concurrent results	4
2.2.1	Additive logistic regression	4
2.2.2	Neural networks	4
2.3	Validation result	5
3	Annealing Data Set	6
3.1	NeuraMorph results	6
3.2	Concurrent results	6
3.2.1	Additive logistic regression	6
3.2.2	Rule based classifier	6
3.3	Validation result	7
4	Arrhythmia Data Set	8
4.1	NeuraMorph results	8
4.2	Concurrent results	8
4.2.1	Additive logistic regression	8
4.2.2	Ensemble classifier, support vector machine and random forest with random sampling	8
4.3	Validation result	8
5	Agaricus Lepiota Data Set	9
5.1	NeuraMorph results	9
5.2	Concurrent results	9
5.2.1	Additive logistic regression	9
5.2.2	Decision tree	9
5.3	Validation result	9

6		10
6.1	NeuraMorph results	10
6.2	Concurrent results	10
6.2.1	Additive logistic regression	10
6.3	Validation result	10

Introduction

This document introduces the validation results of the NeuraMorph algorithm.

Validation has been performed by using NeuraMorph on data sets publicly available on Internet, and comparing the accuracy of prediction made by NeuraMorph with the publicly available results achieved by other algorithms.

To avoid "lucky strike", results introduced here are produced with an instance of the NeuraMorph model selected amongst models produced with a systematic exploration of hyper parameters of the model, and based on its average of prediction accuracy over 5 different random splits of training/test samples.

1 Breast Cancer Wisconsin (Diagnostic) Data Set

Source:

[https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic))

1.1 NeuraMorph results

Training parameters:

Nb inputs/outputs	30/2
Nb samples	569
Percentage of training samples	90%
weakUnitThreshold	0.900000
depth	5
maxLvlDiv	1
nbMaxInputsUnit	2
nbMaxUnitDepth	21
order	2
Time training	17.1s

Results:

Bias prediction (min/avg/sigma/max) and accuracy	c.00 [0.000,0.000,0.000,0.000]	100.00%
	c.01 [0.000,0.000,0.000,0.000]	100.00%
 [0.000,0.000,0.000,0.000]	100.00%
Bias training (min/avg/sigma/max) and accuracy	c.00 [0.000,0.008,0.062,2.000]	99.61%
	c.01 [0.000,0.008,0.062,2.000]	99.61%
 [0.000,0.011,0.062,2.828]	99.61%

1.2 Concurrent results

1.2.1 Multisurface method

"Dr. Wolberg [...] in collaboration with Prof. Mangasarian and two of his graduate students, Rudy Setiono and Kristin Bennett, [developped] a classifier [...] using the multisurface method (MSM) of pattern separation [...] that successfully diagnosed 97% of new cases."

Source:

<http://pages.cs.wisc.edu/~olvi/uwmp/cancer.html>

1.2.2 Additive logistic regression

"J. Friedman, T. Hastie, R. Tibshirani (1998). Additive Logistic Regression: a Statistical View of Boosting. Stanford University."

Predictive accuracy: 96.28%

Source:

<https://www.openml.org/f/77> <https://projecteuclid.org/euclid.aos/1016218223>

1.3 Validation result

NeuraMorph performs better than the concurrent algorithms on this data set.

2 Iris Data Set

Source:

<http://archive.ics.uci.edu/ml/datasets/Iris/>

2.1 NeuraMorph results

Training parameters:

Nb inputs/outputs	4/3
Nb samples	150
Percentage of training samples	90%
weakUnitThreshold	0.975000
depth	2
maxLvlDiv	1
nbMaxInputsUnit	2
nbMaxUnitDepth	11
order	1
Time training	0.1s

Results:

Bias prediction (min/avg/sigma/max) and accuracy	c.00 [0.000,0.000,0.000,0.000]	100.00%
	c.01 [0.000,0.000,0.000,0.000]	100.00%
	c.02 [0.000,0.000,0.000,0.000]	100.00%
 [0.000,0.000,0.000,0.000]	100.00%
Bias training (min/avg/sigma/max) and accuracy	c.00 [0.000,0.000,0.000,0.000]	100.00%
	c.01 [0.000,0.015,0.086,2.000]	99.26%
	c.02 [0.000,0.015,0.086,2.000]	99.26%
 [0.000,0.021,0.000,2.828]	99.26%

2.2 Concurrent results

2.2.1 Additive logistic regression

”J. Friedman, T. Hastie, R. Tibshirani (1998). Additive Logistic Regression: a Statistical View of Boosting. Stanford University.”

Predictive accuracy: 94.67%

Source:

<https://www.openml.org/f/77> <https://projecteuclid.org/euclid.aos/1016218223>

2.2.2 Neural networks

”Classification of Iris data set, Primož Potocnik, Vitaly Borovinskiy”

Predictive accuracy:

Multilayer perceptron: 96.825% Radial basis function network: 96.825% Probabilistic neural network: 95.238%

Source:

http://lab.fs.uni-lj.si/lasin/wp/IMIT_files/neural/doc/seminar8.pdf

2.3 Validation result

NeuraMorph performs better than the concurrent algorithms on this data set.

3 Annealing Data Set

Source:

<https://archive.ics.uci.edu/ml/datasets/Annealing>

3.1 NeuraMorph results

Training parameters:

Nb inputs/outputs	77/5
Nb samples	798
Percentage of training samples	90%
weakUnitThreshold	0.900000
depth	4
maxLvlDiv	4
nbMaxInputsUnit	2
nbMaxUnitDepth	3
order	2
Time training	14.1s

Results:

Bias prediction (min/avg/sigma/max) and accuracy	c.00	[0.000,0.076,0.196,2.000]	96.20%
	c.01	[0.000,0.025,0.113,2.000]	98.73%
	c.02	[0.000,0.000,0.000,0.000]	100.00%
	c.03	[0.000,0.101,0.226,2.000]	94.94%
	c.04	[0.000,0.000,0.000,0.000]	100.00%
	[0.000,0.143,0.196,2.828]	94.94%
Bias training (min/avg/sigma/max) and accuracy	c.00	[0.000,0.050,0.158,2.000]	97.50%
	c.01	[0.000,0.011,0.075,2.000]	99.44%
	c.02	[0.000,0.003,0.037,2.000]	99.86%
	c.03	[0.000,0.036,0.135,2.000]	98.19%
	c.04	[0.000,0.000,0.000,0.000]	100.00%
	[0.000,0.071,0.158,2.828]	97.50%

3.2 Concurrent results

3.2.1 Additive logistic regression

”J. Friedman, T. Hastie, R. Tibshirani (1998). Additive Logistic Regression: a Statistical View of Boosting. Stanford University.”

Predictive accuracy: 99.67%

Source:

<https://www.openml.org/f/77> <https://projecteuclid.org/euclid.aos/1016218223>

3.2.2 Rule based classifier

”Analysis of Different Classification Algorithms Applied to Anneal Dataset Using Data Mining Techniques, V.Kamalakkannan, Dr.D.

Ramyachitra, Computer Science Department, Bharathiar University, Coimbatore, Tamil Nadu, India”

Predictive accuracy:

Naive Bayes: 97.66%

SMO: 97.01%

OneR: 83.65%

J48: 95.88%

Source:

http://www.istpublications.com/temp/16_V_Kamalakkannan_Et_al_.pdf

3.3 Validation result

NeuraMorph performs worst than the concurrent algorithms on this data set.

4 Arrhythmia Data Set

Source:

<https://archive.ics.uci.edu/ml/datasets/arrhythmia>

4.1 NeuraMorph results

4.2 Concurrent results

4.2.1 Additive logistic regression

”J. Friedman, T. Hastie, R. Tibshirani (1998). Additive Logistic Regression: a Statistical View of Boosting. Stanford University.”

Predictive accuracy: 74.12%

Source:

<https://www.openml.org/f/77> <https://projecteuclid.org/euclid.aos/1016218223>

4.2.2 Ensemble classifier, support vector machine and random forest with random sampling

”Arrhythmia Disease Classification and Mobile Based System Design, Soha Samir AbdElMoneem, Hany Hanafy Said, and Amani Anwar Saad, Computer Engineering Department, Arab Academy for Science, Technology and Maritime Transport Abukir, Alexandria, Egypt”

Predictive accuracy: 98.18%

Source:

<https://iopscience.iop.org/article/10.1088/1742-6596/1447/1/012014/pdf>

4.3 Validation result

NeuraMorph performs as well as the concurrent algorithms on this data set.

5 Agaricus Lepiota Data Set

Source:

<https://archive.ics.uci.edu/ml/datasets/mushroom>

5.1 NeuraMorph results

5.2 Concurrent results

5.2.1 Additive logistic regression

”J. Friedman, T. Hastie, R. Tibshirani (1998). Additive Logistic Regression: a Statistical View of Boosting. Stanford University.”

Predictive accuracy: 100.0%

Source:

<https://www.openml.org/f/77> <https://projecteuclid.org/euclid.aos/1016218223>

5.2.2 Decision tree

”Agaricus Lepiota Analysis, Manikanda Prabhu M and Jerin Stanley Daniel J”

Predictive accuracy: 99.0%

Source:

<https://rpubs.com/ManikandaPrabhuM/JerinDaniel-agaricuslepiota>

5.3 Validation result

NeuraMorph performs the concurrent algorithms on this data set.

6

Source:

<https://>

6.1 NeuraMorph results

6.2 Concurrent results

6.2.1 Additive logistic regression

””

Source:

<http://>

6.3 Validation result

NeuraMorph performs as well as the concurrent algorithm(s) on this data set.