International Journal of Scientific Research & Engineering Technology (IJSET)
Vol.13 pp.28-31

# The Improvement Achieved Using BLogReg Feature Selection Algorithm In A Developed Artificial Neural Network Classification

Esra Kaya Ayana [#1], Kenan Morani [*2]

*Control and Automation Department, Yildiz Technical University*
*Istanbul, Turkey*

[1] esrakaya@yildiz.edu.tr
[2] kinan.morani@hotmail.com

***Abstract*— The improvements that can be achieved by using BLogReg feature selection algorithm on the performance of artificial neural network based classifiers is investigated. In that an artificial neural network based classifier is build, trained, and tested using the dataset obtained from mobile robot in machine learning repository (uci). The previous model is then compared to a model applied on a new dataset. The new dataset is obtained by eliminating irrelevant features from the (uci) dataset through the BLogReg algorithm. It is demonstrated that by using the BLogReg algorithm to eliminate unreliable and irrelevant features, both the performance and the accuracy of the classifier can be sufficiently increased.**

***Keywords*— Mobile robot, Artificial Neural Network, Feature Selection, BLogReg Algorithm, Bayesian Method, Spare Logistic Regression, Mean Square Error.**

## I. INTRODUCTION

In order to improve the mobile robot applications several studies have been conducted on Scitos G5, a mobile robot platform. However, not enough investigate has been done on the classification and optimum accuracy of classification [1]. Furthermore, few studies have dealt with feature selection implementations on classification models, some of those include areas such as robotic, social and health [2]. Yet little emphasize have been given to the BLogReg algorithm which is related to Bayesian method [3].

Classification studies are vitally important in mobile robots for determining direction of the robot’s movements and applying feature selection to the classification is essential to improve the models [4,5].

This research was conducted on a dataset from mobile robots. Developed artificial neural network model have been used to determine the directions of the robot. Two similar types of ANNs models have been implemented. One type was implemented on the given dataset and a similar model was implemented on a new dataset obtained from the given dataset after eliminating irrelevant features using BLogReg feature selection algorithm [6,7,8]. The two models were then compared in terms of performance and accuracy. The study illustrates the improvements in both the model’s accuracy and performance of the classifier when the feature selection algorithm is implemented on the dataset.

## II. BAYESIAN, SLOGREG AND BLOGREG

The feature algorithm chosen for this study is Bayesian Logistic Regression (BLogReg). This algorithm has been mainly preferred because it is much faster and it is free from any selection bias which fits with our dataset where all our input sensors are of equal importance [9,10].

Bayesian Equation negative log-likelihood can be described as follows:

$$E_{\vartheta} = \sum_{I=1}^{\ell} g\left\{-y_i f\left(x_i\right)\right\} \tag{1}$$

$\vartheta$ is assumed to be an independent and identically distributed sample from Bernoulli distribution.

$\ell$ is the training examples

$x_i$ is a vector of measurements for the i-th example

$y_i$ is the class that *i-th* example belongs to given that:

$$g\left\{\xi\right\} = \log\left(1 + e^{\left(\xi\right)}\right)$$

Spares Logistic Regression equation can be obtained from Equation (1) by adding a regularization term as follows:

$$M = E_{\vartheta} + \lambda E_{\alpha} \tag{2}$$

Where

$\lambda$ is a regularization parameter to control the bias (strictly positive)

$\alpha$ is the bias parameter

Where

$$E_{\alpha} = \sum_{i=1}^{d} \left|\alpha_i\right|$$

By eliminating the $\lambda$ parameter from Equation (2), a revised optimization criterion can be obtained called Bayesian Logistic Regression (BLogReg).


ISSN 1737-9296

### A. *Materials and Methods*

Scitos G5 mobile robot was used to navigate some environment. The robot uses the data coming from 24 ultrasonic sensors as inputs and provides 4 classified outputs to guide the robot while navigating (moving slightly right, sharply right, forward, or slightly left). To help the robot navigate through processing the input data, an artificial neural network was developed. Table 1 shows part of the input data taking from 5456 samples.

TABLE I
INPUT DATA FROM CHOSEN SENSORS

| Sensor 1 | Sensor 2 | Sensor 12 | Sensor 20 | Sensor 24 |
|---|---|---|---|---|
| 0.438 | 0.498 | 1.687 | 0.445 | 0.429 |
| 0.438 | 0.498 | 1.687 | 0.449 | 0.429 |
| 0.438 | 0.498 | 1.687 | 0.449 | 0.429 |
| 0.437 | 0.501 | 1.687 | 0.449 | 0.429 |
| 0.438 | 0.498 | 1.687 | 0.449 | 0.429 |
| 0.439 | 0.498 | 1.686 | 0.446 | 0.43 |
| 0.44 | 5 | 1.684 | 0.451 | 0.432 |
| 0.444 | 5.021 | 1.68 | 0.453 | 0.436 |
| 0.451 | 5.025 | 1.673 | 0.457 | 0.442 |
| 0.458 | 5.022 | 1.666 | 0.462 | 0.449 |
| 0.465 | 0.525 | 1.658 | 0.467 | 0.457 |
| 0.473 | 0.533 | 1.651 | 0.469 | 0.465 |

Before modelling the dataset was divided into a training set, which takes 60% of the total data or 3275 samples, and test set, which takes 40% of the total data or 2181 samples. Fig.1 shows the neural network module with 40 hidden-layers.

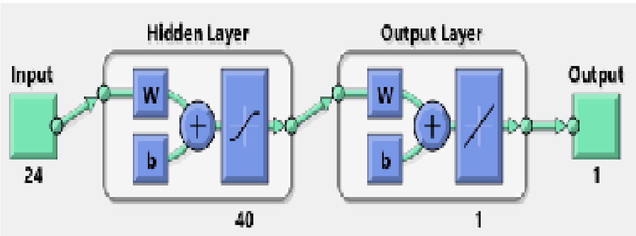


Fig. 1 The artificial neural network module

## III. THE EFFECT OF BLOGREG ALGORITHM ON CLASSIFICATION PROBLEM

### A. *The robot and the dataset*

The SCİTOS G5 belongs to the metraLab in Germany (http://metralabs.com). The robot is designed for indoor uses. The weight of the robot is 60 kilograms and the maximum speed is 1.4 meters per second. The robot is capable of rotating 360 degrees. Fig. 2 shows the Scitos G5 robot.

The data were collected by sampling at a rate of 9 samples per seconds generating a database of 5456 examples. This data are then used to train the neural network. In order for the robot to make directional decisions it follows IF-THEN algorithm measuring front-distance and left-distance, while also calculating distances to the right and behind the algorithm goes as follows [11,12]

```
  Wall-Following Algorithm
if leftDist > 0.9
  then
  if frontDist <= 0.9
    then Stop and turn to the right
  else Slow down and turn to the left
        if frontDist <= 0.9
          then Stop and turn to the right
        else if leftDist < 0.55
          then Slow down and turn to the right
        else Move forward
```

Fig. 2 The Scitos G5 mobile robot

### B. *Applying BLogReg Algorithm to Eliminate Features*

Using the BLogReg algorithm on the given dataset eliminated some input sensors. The more significant sensors' inputs according to the algorithm are listed in Table 2 along with their weights.

TABLE III
THE MORE SIGNIFICANT INPUTS AND THEIR WEIGHTS

| The More Significant Input Sensors | Weights of the Sensors |
|---|---|
| 24 | 10.968690 |
| 11 | 10.202424 |
| 15 | 8.601046 |
| 18 | 8.168761 |
| 17 | 7.731453 |
| 13 | 7.242650 |
| 9 | 5.566802 |
| 23 | 3.972673 |
| 1 | 3.031587 |
| 4 | 2.919844 |
| 5 | 2.762847 |
| 8 | 2.626482 |
| 16 | 2.468423 |
| 3 | 2.122611 |

With the inputs and weight values in Table 2, only the first six sensors were chosen to be the inputs to the improved neural network.


ISSN 1737-9296

### *C. Comparing the Performance and Accuracy*

To compare the performance and the accuracy of the Artificial neural networks on the original dataset and on the dataset obtained after applying BLogReg algorithm, the MSE and the regression were potted and compared. Fig. 3 shows the performance of the module applied to the original dataset through MSE and Fig. 4 shows the performance of the module on the derived dataset.

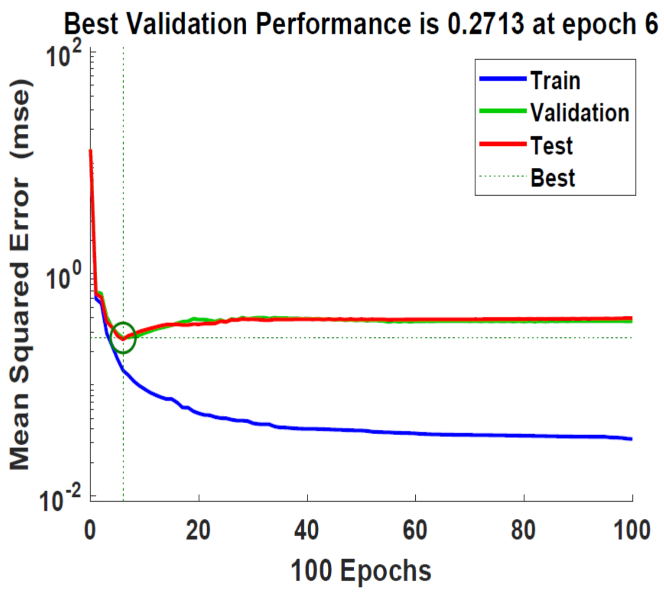


Fig. 3 The performance on the original dataset

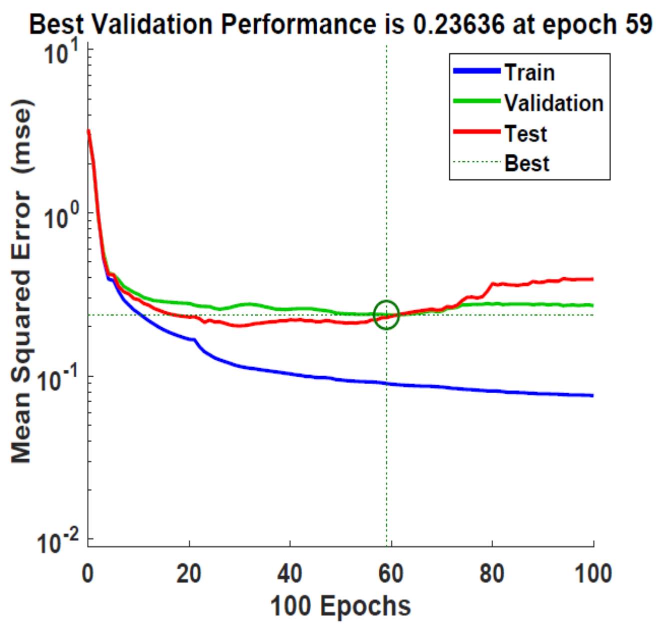


Fig. 4 The performance on the derived dataset

Fig. 4 shows best validation performance of 0.23636 for the model, and in Fig. 3 the best validation performance is 0.2713.

The accuracy of both modules was compared through the regression curve. Fig. 5 shows the regression curve for the module on the original dataset and Fig. 6 shows the regression for the module applied on the derived dataset.

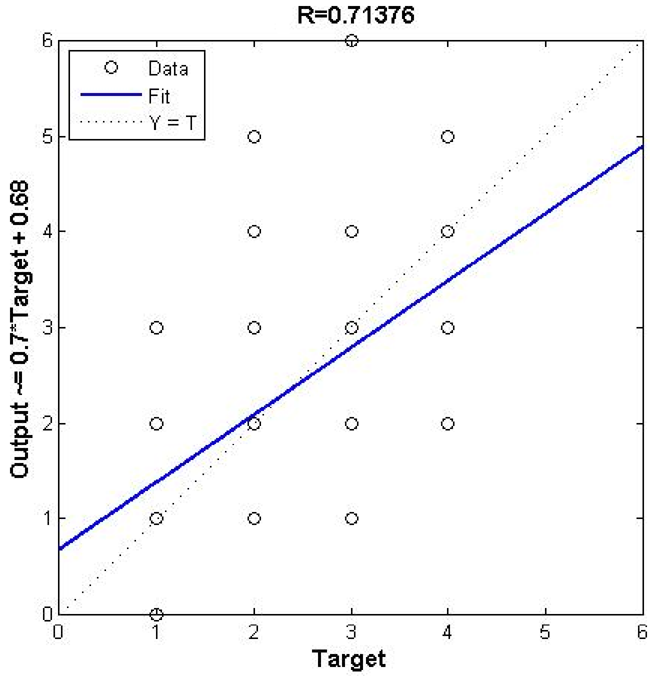


Fig. 5 The accuracy on the original dataset

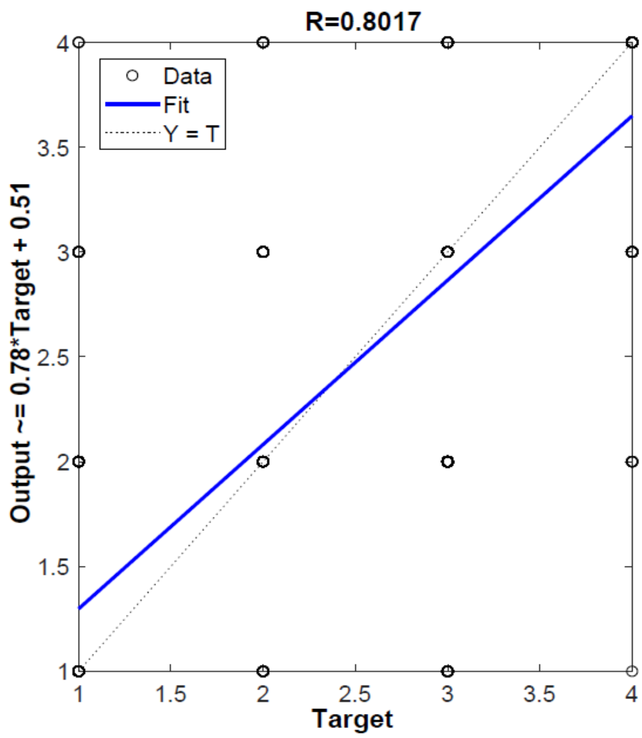


Fig. 6 The accuracy on the derived dataset


ISSN 1737-9296

The regression line in Fig. 6 shows a linear model regression $Y = (0.78 * X + 0.5)$ with R-squared being about 0.642 and thus about 64% of the variability can be explained through model. In Fig. 5 R-squared is approximately 0.509 with about 51% of the variability explained through the model.

From the previous figures and through MatLab, the performance of the module is seen to increase from 0.6868 to 0.8326 and the regression value increases from 0.71376 to 0.8017 indicating more accuracy in the module [13].

## IV. Conclusion

The work shows that improvement in classification problems can be made using BLogReg algorithm to help robots navigate their environments. The improvement is in achieving more accuracy in the module, minimum error and faster implementation.

The paper shows the improvement that can be made using the BLogReg algorithm to eliminate insignificant inputs, where all inputs are identical sensors that have the same contribution to the decision making of direction.

Further work can be carried out; one could use the BLogReg algorithm on different datasets, where inputs are not of equal importance. At the same time, other algorithms could be trialed and tested in the field of robotic navigation or other related fields. It is also important to try the BLogReg algorithms on bigger or even smaller datasets to further verify the results achieved in this work.

## Acknowledgment

The work has been conducted at the Control and Automation Engineering Department in the Electrical and Electronic Faculty at Yildiz Technical University. The authors are grateful to the facilities found in the department and all the help they have received from the existing academicians as well as the staff.

## References

[1] E. Kaya Ayana, "Öznitelik seçiminin geliştirilen bir yapay sinir ağı sınıflandırılması üzerine etkisi," Master thesis, Turkish Institute of Science, Istanbul, Turkey, Jan. 2016.

[2] K. Morani, E. Kaya Ayana, and Ş.N. Engin, "Developement of prediction in clients' consent to a bank term deposit using feature selection," *2018 6th International Conference on Control Engineering & Information Technology (CEIT), IEEE, 2018*, p.1-5.

[3] S. Haykin, Neural Networks and Learning Machines, McMaster University, Hamilton, Ontario, Canada, 2001, third edition.

[4] M.S. Esfahania, J. Knighta, A. Zollanvaria, B.J. Yoona, and E.R. Dougherty, "Classifier design given an uncertainty class of feature distributions via regularized maximum likelihood and the incorporation of biological pathway knowledge in steady-state phenotype classification," *Pattern Recognition*, vol.46, pp.2783–2797, Oct. 2013.

[5] T. Koskela, M. Lehtokangas, J. Saarinen, and K. Kaski, "Time series prediction with multilayer perceptron, FIR and Elman neural networks," *in Proceedings of the World Congress on Neural Networks, INNS Press San Diego, USA, 1996*, p. 491-496.

[6] E.R. Dougherty, J. Hua, and C. Sima, "Performance of feature selection methods," *Current Genomics*, vol.10, pp.365-374, Sept. 2009.

[7] G.C. Cawley, and N.L.C. Talbot, "Gene selection in cancer classification using sparse logistic regression with bayesian regularization," *Bioinformatics*, vol.1, pp.2348-55, 2006.

[8] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu. (2019) feature selection. [Online]. Available: http://featureselection.asu.edu/algorithms.php

[9] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu, "Feature selection: A data perspective," *ACM Computing Surveys*, vol. 50(6), Jan. 2018.

[10] V. Klingspor, K.J. Mork and A.D. Rieger. (1995) learning concepts from sensor data of a mobile robot; set of datasets. [Online]. Available: https://archive.ics.uci.edu/ml/dataset/Mobile+Robots

[11] G. Y. Wong, and W. M. Mason, "The hierarchical logistic regression model for multilevel analysis," *Journal of the American Statistical Association*, vol. 80, pp. 513-524, 1985.

[12] A.L Freire, A.E. Barreto, M. Veloso, and A.T. Varel, "Short-term memory mechanisms in neural network learning of robot navigation tasks:a case study," *2009 6th Latin American Robotics Symposium (LARS 2009), IEE*, 2009, p.1-6.

[13] W. H. Ibrahim, A.A.A Osman, and Y.I. Mohamed, "MRI brain image classification using neural networks," *2013 International Conference on Computing, Electrical and Electronic Engineering (ICCEEE). IEEE, 2013*, p.253-258.


ISSN 1737-9296