# Heterogeneous Processing for Estimating the Parameters of a Binary Logistic Regression Model Using the Fisher Scoring Algorithm

Castillo-Méndez Luis Eduardo, Vera-Parra Nelson Enrique, Medina-Daza Rubén Javier

*Abstract*—**This article presents the implementation and evaluation of a heterogeneous processing model (multi-core/many-core) to estimate the parameters of a binary logistic regression model using the Fisher scoring algorithm. The model was implemented on a heterogeneous CPU/GPU platform employing CUDA and was evaluated with five datasets of different sizes (ranging from 10000 to 30000 records). The evaluation results showed a speed-up of up to 6.09X without affecting the quality of the estimate when compared to sequential implementation. The implementation is available in the following repository: https://github.com/Parall-UD/ParallelFischerScoring_binomial.**

*Index Terms*—**Binary logistic regression model, Fisher scoring algorithm, heterogeneous computing, parallel computing.**

## I. Introduction

LOGISTIC regression is a statistical model for estimating the probability of a categorical variable, that is, a variable that can take a limited number of categorical values, which can be explained by a set of variables called independent or predictor variables. The model is called binary logistic regression when the categorical variable is a dichotomous variable, i.e. it has only two categories [1].

Logistic regressions are generally implemented as part of the procedure and in the final part of different machine learning algorithms applied in classification tasks; for example, it is very common to find activation functions based on logistic regression in the output layers of convolutional neural networks [2]-[4].

Logistic regression analysis can be considered a special case in the theory of generalized linear models, so the estimation of its parameters can be conducted with a classical approach as well as with the employment of maximum likelihood estimation by using the Fisher scoring algorithm [5]; this algorithm is an adaptation of the Newton-Raphson method [6], which has greater robustness to poor initial values [7] and advantages in terms of convergence, especially due to its low dependence on specific data values [8].

Manuscript received May 7, 2021; revised September 10, 2021. This work was fully supported by the Universidad Distrital Francisco José de Caldas, Bogotá – Colombia.

Castillo-Méndez Luis Eduardo is a PhD candidate of Universidad Nacional, Bogotá – Colombia. He is the leader of the GEFEM research group, Universidad Distrital Francisco José de Caldas, Bogotá – Colombia, phone: +57-3002011213; (e-mail: lecastillom@udistrital.edu.co).

Vera-Parra Nelson Enrique is a professor of electronic engineering program and coordinator of the master’s program in information and communication sciences at the Universidad Distrital Francisco José de Caldas, Bogotá – Colombia, phone: +57-3012401090; (e-mail: neverap@udistrital.edu.co).

Medina-Daza Rubén Javier is a professor of cadastral engineering and PhD in Engineering programs at the Universidad Distrital Francisco José de Caldas, Bogotá – Colombia, phone: +57-3208998714; (e-mail: rmedina@udistrital.edu.co).

## II. Binary Logistic Regression

Let $Y$ be a binary dependent variable which can take two possible values, namely 0, 1. Consider observations $y_1, y_2, y_3, ..., y_p$ of Y and let $X_1, X_2, X_3, ..., X_p$, be the explanatory variables with assigned observations $x_{i1}, x_{i2}, x_{i3}, ..., x_{ip}$ associated to each $y_i$; also consider that $x_{i1}$ usually takes the value of 1. Then, the binary logistic regression for Y can be expressed as follows:

$$\pi_i = P\left[Y_i = 1/ x_{i1}, x_{i2}, x_{i3}, ..., + x_{ip}\right] = \frac{1}{1 + e^{-\sum_{j=1}^{p} \beta_j x_{ij}}} \tag{1}$$

where $\beta_j$ are the model parameters with $i = 1, 2, 3, ...n$ and $j = 1, 2, 3, ...p$. The purpose of the regression is to estimate the model parameters that best fit the functional expression of the model.

## III. Generalized Linear Models

The referential framework of this proposal is inspired by the theory and methodologies on generalized linear models, developed by Nelder [9], McCullagh [10], Cepeda [11], and Dobson [12].

Let $Y$ be a variable called dependent variable, or explained variable, with observations $y_1, y_2, y_3, ..., y_p$, and let variables $X_1, X_2, X_3, ..., X_p$ be explanatory variables. Observations $x_{i1}, x_{i2}, x_{i3}, ..., x_{ip}$ are associated to each observation $y_i$, with $x_{i1}$ usually being equal to 1. Therefore, a generalized linear model for $Y$, explained by $X_1, X_2, X_3, ..., X_p$ is said to be a model consisting of three components: the random component, the systematic component, and the link function component.

(i) The random component identifies the dependent variable $Y$ and its probability distribution. The probability density function for any $y$ of Y is as follows:

$$f(x) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y;\phi)\right\} \quad (2)$$

where $\mu = E[y] = b'(\theta)$, $Var[y] = a(\phi)b''(\theta)$, and $a(\phi)$ is called dispersion function. Parameter $\theta$ is called canonical parameter.

(ii) The systematic component specifies the explanatory variables through the following relationship:

$$\eta = \mathbf{X}\boldsymbol{\beta} \cdot\cdot \quad (3)$$

where,

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, ..., \beta_1)^t \quad (4)$$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{1p} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{np} \end{bmatrix} \quad (5)$$

$$\eta = (\eta_1, \eta_2, \eta_3, ..., \eta_n)^t \quad (6)$$

$$\eta_i = \sum_{j=1}^{p} \beta_j x_{ij} \quad (7)$$

$\eta$ is known as a linear predictor.

(iii) The link function component is a function $g$, of the expected value of the dependent variable that is equated to the linear predictor, i.e. $g(\mu) = \eta$; in other words, for each $y_i$ as a random variable,

$$g_i(\mu_i) = \sum_{j=1}^{p} \beta_j x_{ij} \quad (8)$$

The generalized linear model is, therefore, as follows:

$$\mu_i = E[Y_i] = g^{-1}\left(\sum_{j=1}^{p} \beta_j x_{ij}\right) \quad (9)$$

## IV. BINARY LOGISTIC MODEL AS A GENERALIZED LINEAR MODEL

The binary logistic model is a particular case of a generalized linear model, that is: if $Y$ is a random Bernoulli-type variable with parameters $\pi$ and $Y = \{0,1\}$, the associated probability function belongs to the exponential family, since

$$\begin{aligned} p(y) &= \pi^y (1-\pi)^{1-y} \\ &= \exp\{y\ln(\pi) + (1-y)\ln(1-y)\} \\ &= \exp\{y\ln(\pi) - \ln(1-y) + \ln(1-\pi)\} \\ &= \exp\{y\theta + \ln(1+e^{\theta})\} \end{aligned} \quad (10)$$

where $\theta = \ln\left[\frac{\pi}{1-\pi}\right]$, $b(\theta) = \ln\left[1-e^{\theta}\right]$, $a(\phi) = 1$, and $c(y;\phi) = 0$. Subsequently,

$$\ln\left[\frac{\pi_i}{1-\pi_i}\right] = \sum_{j=1}^{p} \beta_j x_{ij} \quad (11)$$

The following binary logistic model is obtained by solving for $\pi_i$:

$$\pi_i = \frac{1}{1+e^{-\sum_{j=1}^{p} \beta_j x_{ij}}} \quad (12)$$

## V. CLASSIC PARAMETER ESTIMATION IN A GENERALIZED LINEAR MODEL

Based on generalized linear models, the value of the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$, can be determined by maximizing the likelihood function $\iota(\beta) = \sum_{i-1}^{n} \iota_i(\boldsymbol{\beta}) = \sum_{i=1}^{n}\left\{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i;\phi)\right\}$ using the Newton-Raphson method. The resulting likelihood equations are as follows:

$$\sum_{i=1}^{n} \frac{(y_i - \mu_i)x_{ij}}{Var[Y_i]} \frac{\partial \mu_i}{\partial \eta_i} = 0 \quad (13)$$

with $j = 1, 2, 3..., p$. If

$$\mathbf{q} = \left(\frac{\partial \iota(\boldsymbol{\beta})}{\partial \beta_1}, \frac{\partial \iota(\boldsymbol{\beta})}{\partial \beta_2}, \frac{\partial \iota(\boldsymbol{\beta})}{\partial \beta_3}, ..., \frac{\partial \iota(\boldsymbol{\beta})}{\partial \beta_p}\right)^t \quad (14)$$

and $H$ is the Hessian matrix such that:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_1^2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_3} & \cdots & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_P} \\ \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_2^2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_1 \partial \beta_2} & \cdots & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_2 \partial \beta_p} \\ \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_3 \partial \beta_2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_3 \partial \beta_2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_3^2} & \cdots & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_3 \partial \beta_P} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_P \partial \beta_2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta P \partial \beta_2} & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_P \partial \beta_3} & \cdots & \frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial \beta_P^2} \end{bmatrix} \quad (15)$$

Then, in using the maximum likelihood estimation of $\widehat{\boldsymbol{\beta}}$, applying the Newton-Raphson method, the $(k+1)$ iteration will be as follows:

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} - (H^{(k)})^{-1}\mathbf{q}^{(k)} \tag{16}$$

The Fisher scoring algorithm is proposed as a variant of the Newton-Rapshon method; this algorithm uses $-E\left[\frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial\beta_a\partial\beta_b}\right]$ instead of $\frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial\beta_a\partial\beta_b}$ from the $(k+1)$ Newton-Raphson iteration onwards, where $a$ and $b$ are integer values ranging from 1 to $p$. In other words, the proposal involves replacing the $-H$ matrix of the Newton-Raphson method with the $I$ information matrix. Explicitly, the procedure is as follows. Given that,

$$-E\left[\frac{\partial^2 \iota(\boldsymbol{\beta})}{\partial\beta_a\partial\beta_b}\right] = \sum_{i=1}^{n}\frac{x_{ia}}{Var[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 \tag{17}$$

$$I = \mathbf{X}^t W\mathbf{X} \tag{18}$$

is defined, where

$$W = diag\left\{\frac{1}{Var[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 : i = 1, 2, \ldots n\right\} \tag{19}$$

The $(k+1)$ iteration to obtain $\hat{\boldsymbol{\beta}}$ is as follows:

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} + (I^{(k)})^{-1}\boldsymbol{q}^{(k)} \tag{20}$$

which is the same as

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} + (I^{(k)})^{-1} + \mathbf{q}^{(k)} \tag{21}$$

Moreover, this last expression is equivalent to:

$$(X^t W^{(k)}\mathbf{X})\boldsymbol{\beta}^{(k+1)} = (X^t W^{(k)}\mathbf{X})\boldsymbol{\beta}^{(k)} + \mathbf{q}^{(k)} \tag{22}$$

which is the same as

$$\boldsymbol{\beta}^{(k+1)} = (\mathbf{X}^t W^{(k)}\mathbf{X})^{-1}\mathbf{X}^t W^{(k)} Z^{(k)} \tag{23}$$

where $Z^{(K)}$ is a matrix representing the $(k)-th$ iteration of matrix Z. Matrix Z is of order $n \times 1$ such that,

$$z_i = \eta_i + (y_i - \mu_i)\frac{\partial\mu_i}{\partial\eta_i} \tag{24}$$

which, in matrix terms, is the same as:

$$Z = X\beta + W^{-1}(Y - \mu) \tag{25}$$

Considering the binary logistic model as a special case of the generalized linear model, derived from the Bernoulli probability distribution with parameter $\pi$, for this case:

$$W = diag\{\pi_i(1-\pi_i) : i = 1, 2, , \ldots, n\} \tag{26}$$

## VI. Heterogeneous Processing Model for Implementing the Fisher Scoring Algorithm

A heterogeneous computing platform is defined as a system comprising at least two different types of processors for incorporating specialized processing capabilities to perform particular tasks [13], [14]. A heterogeneous system typically consists of one or more CPUs acting as the main processing unit (usually called host) and one or more other processing devices, such as GPUs (Graphics Processing Units), DSPs (Digital Signal Processors), and FPGAs (Field Programmable Gate Arrays), acting as accelerators. It is also possible to find the integration of two or more types of processors on a single chip; for example, an APU (Accelerated Processing Unit) is a microprocessor that integrates a multi-core CPU and a GPU by means of a high-speed bus.

The efficient use of these heterogeneous platforms requires the adaptation of the algorithms to a processing model that allows the best use of each architecture. In this work, the Fisher scoring algorithm was adapted to a heterogeneous processing model based on convenient

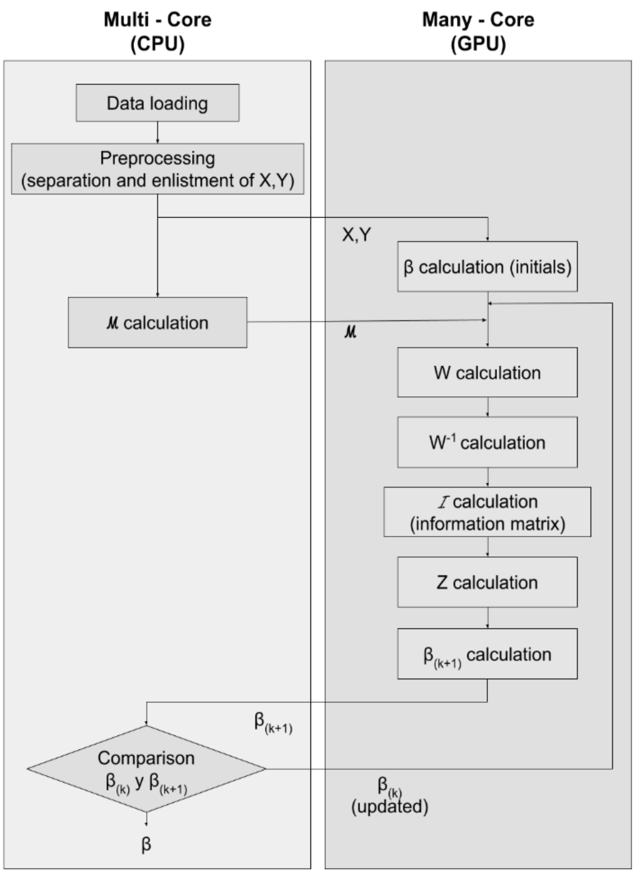


Fig. 1. Schematic heterogeneous processing model for implementing Fisher scoring algorithm.

segmentation and distribution of tasks between multi-core and many-core architectures (see fig. 1). The convenience of the sectioning processes and the way processes are served by GPU was based on three criteria: (1) the computational requirement of the process, (2) the nature of the process (which could be adaptable to massive parallel processing), and (3) the impact in terms of data transfer (a high transfer of data between architectures may not be required).

## VII. RESULTS AND DISCUSSION

The heterogeneous processing model (multi-core/many-core) for estimating the parameters of a binary logistic regression model through the Fisher scoring algorithm was implemented by using Python and CUDA for a CPU/GPU-type platform. The implementation was evaluated by simulating a binary logistic regression model with two explanatory variables, that is, a model of the form $Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2$, $P(Y) = 1/(1 + e^{-Y})$, which yields the estimators of parameters $\beta_S$.

The evaluation was twofold. A computational performance evaluation was first conducted to measure the speed-up gain obtained after applying the heterogeneous processing model (CPU/GPU vs. CPU) to the Fisher scoring algorithm. Evaluation of the parameter estimation error was also conducted (the error of $\beta_S$ estimators) by comparing the predictions of the binary logistic regression model with the observed data and computing the corresponding Root Mean Square Error (RMSE).

Computational performance evaluation consisted in measuring the time required by the heterogeneous implementation (CPU/GPU) of the Fisher algorithm, when computing the linear regression, and compare it to the time spent by the homogeneous (CPU) counterpart. This comparison was conducted for 5 datasets involving 10000, 15000, 20000, 25000 and 30000 entries. Table 1 shows that the speed-up gains lie between 2X a 6X, which evinces the improved performance obtained from the heterogeneous processing proposal with its corresponding parallelization patterns. Moreover, an almost-linear relation can be observed between the speed-up gain and the size of the dataset, indicating that the heterogeneous processing proposal achieves more significant improvements as the size of the dataset grows.

For the evaluation of errors in the estimation of the regression coefficients, RMSE was computed for both the predictions of the model using the values of $\beta_S$ obtained with the Fisher scoring algorithm and the predictions

TABLE I
PROCESSING TIMES AND SPEED-UP

| Size(records) | Time(s)-CPU | Time(s)-CPU/GPU | Speed-up |
|---|---|---|---|
| **10000** | 11.89 | 4.88 | 2.23x |
| **15000** | 28.34 | 7.65 | 3.70x |
| **20000** | 41.71 | 7.89 | 5.28x |
| **25000** | 85.78 | 15.22 | 5.63x |
| **30000** | 115.8 | 18.99 | 6.09x |

TABLE II
OBSERVED MODEL VS. ESTIMATED MODEL
DATA SET SIZE = 1000 RECORDS

| coefficients | Observed coefficients | | | Estimated coefficients | | | Model Global RMSE |
|---|---|---|---|---|---|---|---|
| | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | |
| s+ | 0,1 | 0,9 | 1,2 | 0,0996 | 0,7914 | 1,0655 | 0,020 |
| l+ | 5 | 6 | 4 | 4,8169 | 5,2559 | 3,347 | 0,050 |
| sl+ | 0,3 | 0,1 | 4 | 0,1359 | 0,6047 | 4,3672 | 0,062 |
| s- | -0,2 | -0,6 | -1,5 | -0,0373 | -0,619 | -1,4539 | 0,030 |
| l- | -2 | -3 | -4,5 | -3,1226 | -4,1853 | -6,3169 | 0,045 |
| sl- | -0,2 | -3 | -0,4 | -0,2433 | -3,4479 | -0,4089 | 0,022 |
| s+- | 0,2 | -0,7 | -1,5 | -0,012 | -0,7176 | -1,7855 | 0,045 |
| l+- | 2 | -3 | 4,5 | 1,834 | -3,1106 | 5,9161 | 0,078 |
| sl+- | -0,2 | 3 | 0,4 | -0,3889 | 2,2636 | 0,7393 | 0,089 |

$\beta_S$ -> s: small; l: large; +: positive; -: negative

TABLE III
OBSERVED MODEL VS. ESTIMATED MODEL
DATA SET SIZE = 20000 RECORDS

| coefficients | Real coefficients | | | Estimated coefficients | | | Model Global RMSE |
|---|---|---|---|---|---|---|---|
| | f1 | f2 | f3 | f1 | f2 | f3 | |
| s+ | 0,15 | 0,5 | 0,1 | 0,1333 | 0,5042 | 0,1145 | 0,0051 |
| b+ | 4 | 3 | 2 | 3,2583 | 2,1177 | 1,3994 | 0,0052 |
| sb+ | 4 | 1 | 1,5 | 3,9412 | 0,9977 | 1,3485 | 0,0126 |
| s- | -1 | -0,8 | -0,7 | -1,027 | -0,821 | -0,709 | 0,0054 |
| b- | -5 | -3 | -9 | -4,616 | -2,29 | -8,4886 | 0,0044 |
| sb- | -3 | -0,8 | -1,5 | -2,982 | -0,749 | -1,4952 | 0,0045 |
| s+- | -1 | 1 | -1 | -0,9892 | 1,0254 | -1,0067 | 0,0052 |
| b+- | -5 | 3 | -8 | -5,1719 | 3,054 | -8,4938 | 0,0139 |
| sb+- | -3 | 0,8 | 1,5 | -2,9958 | 0,856 | 1,4656 | 0,0072 |

$\beta_S$ -> s: small; l: large; +: positive; -: negative

obtained from other 9 logistic regression models. The 9 models used for comparison involved a variety of coefficient types, e.g. positive and negative coefficients as well as small-valued (between 0 and 1) and large-valued (between 1 and 10) coefficients; also, some of the models employed combinations of coefficient types. For this error evaluations, the models were applied to two datasets of different sizes, namely datasets with 1000 and 20000 entries. It was observed that the estimated values of the parameters remained unaltered when implementing the heterogeneous processing proposal based on the Fisher scoring. This observation indicates that the use of large-scale parallel computation does not affect the values obtained from the algorithm. Tables II and III show values of RMSE ranging from 0.005 to 0.089, which support the choice of the Fisher scoring algorithm as a convenient solution, along with heterogeneous computation, to the problem of parameter estimation in binary linear regression models. The results indicate that the proposed solution is convenient in terms of both computational performance and parameter estimation quality. Moreover, Tables II and III

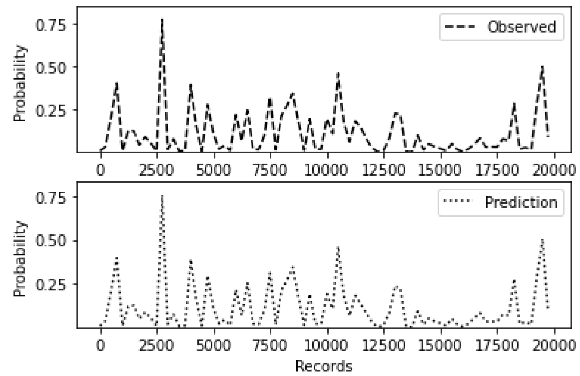


Fig. 2. Probability predicted vs probability observed. Best case ($\beta_1$=-3, $\beta_2$=-0.8, $\beta_3$=-1.5; size data set = 20000; RMSE=0.005,)

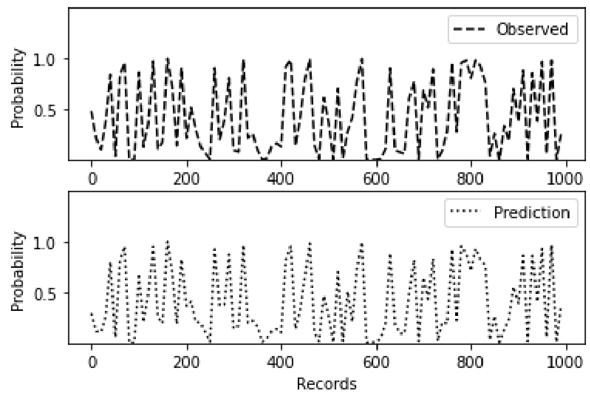


Fig. 3. Probability predicted vs probability observed. Worst case ($\beta_1$=-0.2, $\beta_2$=3, $\beta_3$=0.4; size data set = 1000; RMSE=0.089)

show that the values of RMSE have an inverse relation with the size of the dataset, while a direct relation can be observed with the values of $\beta_S$; however, there is no apparent relation between the values of RMSE and the sign of the values of $\beta_S$.

Figures 2 and 3 show the values of probability predicted by the model compared to the empirical probability of the observations for two regression models, namely the models with the best and worst values of RMSE, respectively.

## VIII. CONCLUSIONS

The estimation of the parameters of a binary logistic regression model, employing the Fisher scoring algorithm, is a process with a computation time depending exponentially on the size of the input data set. This is a significant obstacle for its implementation since regression problems usually require large data sets. This project has shown that heterogeneous computing represents a real and efficient solution to such an exponential dependence. The proper integration of many-core elements (based on massive parallelization), which shows better performance as the data set increases, allows this dependence to have a favorable linear behavior.

The heterogeneous processing model described in this article allows an efficient implementation of the Fisher scoring method with large data sets, without affecting the precision of the estimated $\beta_s$, when compared to a homogeneous sequential implementation (CPU).

## ACKNOWLEDGMENT

The authors would like to thank Universidad Distrital Francisco José de Caldas and the GICOGE and GEFEM research groups for supporting the development and testing of the project.

## REFERENCES

[1] Cox, D. R., The regression analysis of binary sequences, journal of the royal statistical society: Series B (Methodological), 20(2), 215-232 (1958).

[2] Gulli, A. and Pal, S., Deep learning with Keras, Packt publishing Ltd, India (1958).

[3] Muhammad Aqeel Aslam, Cuili Xue, Manhua Liu, Kan Wang, and Daxiang Cui, "Classification and prediction of gastric cancer from saliva diagnosis using artificial neural network," Engineering letters, vol. 29, no.1, pp10-24, 2021

[4] Kazuya Kishida, Kei Hasegawa, and Kiyotaka Kamata, "Construction of classifier of myoelectric signals by using ANNs," Lecture notes in engineering and computer science: Proceedings of the world congress on engineering and computer science 2019, 22-24 october, 2019, San Francisco, USA, pp356-361

[5] Ypma, T. J., Historical development of the Newton–Raphson method, SIAM review, 37(4), 531-551 (1995).

[6] Longford, N. T., A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects, Biometrika, 74(4), 817-827 (1987).

[7] Jennrich, R. I. and Sampson, P. F., Newton-Raphson, and related algorithms for maximum likelihood variance component estimation. Technometrics, 18(1), 11-17 (1976).

[8] Schworer, A. and Hovey, P., Newton-Raphson versus Fisher scoring algorithms in calculating maximum likelihood estimates, mathematics day - Electronic proceeding, University of Dayton, United States (2004).

[9] Nelder, J. A., W. W. M. , Generalized linear models, journal of the royal statistical society, 3, 370 - 384 135 (1972).

[10] McCullagh, J., N. J. , Generalized linear models, Chapman - Hall, 3, London (1972).

[11] Cepeda, E., Modelagem da variabilidade em modelos lineares generalizados, doctoral thesis, Universidade Federal do Rio de Janeiro, Brasil (2001).

[12] Dobson, A. J., An introduction to generalized linear models, doctoral thesis, CRC Press (2010).

[13] Amar Shan, Heterogeneous processing: a strategy for augmenting Moore's law, Linux J., 142, 1-7 (2006).

[14] Hejun Xuan, Shiwei Wei, Xuelin Zhao, Yang Zhou, Xingpo Ma, Daohua Liu, and Yanling Li, "Unavailable time aware scheduling of hybrid task on heterogeneous distributed system," IAENG International Journal of Applied Mathematics, vol. 50, no.1, pp133-146, 2020

**Castillo-Méndez Luis Eduardo.** He is a PhD candidate and magister in statistics from the Universidad Nacional de Colombia. Professor and researcher at the Universidad Distrital Francisco José de Caldas (Bogotá, Colombia). Director of the GEFEM research group.

**Vera-Parra Nelson Enrique.** Professor and coordinator of the master's program in information and communication sciences at the Universidad Distrital Francisco José de Caldas (Bogotá, Colombia), Doctor of engineering from the same University, electronic engineer from the Universidad Surcolombiana (Neiva, Colombia), Researcher in parallel computing, high performance computing, science data and bioinformatics.

**Medina-Daza Rubén Javier.** Ph.D. in computer science, emphasis on geographic information systems, Universidad Pontificia de Salamanca Madrid-España campus. accredited in research proficiency – Diploma Advanced Studies (DEA) – in geographic information systems, Universidad Pontificia de Salamanca Madrid-España campus. Master in teleinformatics, specialist in software engineering, Specialist in geographic information systems and bachelor of mathematics from the Universidad Distrital Francisco José de Caldas (Bogotá, Colombia). Full professor time complete assigned to the faculty of engineering from the same University, teacher of the curriculum Project cadastral engineering and geodesy, master's in information sciences and communications and the doctorate of engineering of the faculty of engineering of the same University.