ARTIFICIAL NEURAL NETWORKS – A SURVEY ABOUT HARDWARE AND SOFTWARE USE

Darío Baptista, Filipe Freitas, Rita Vasconcelos and Fernando Morgado-Dias

Madeira Interactive Technologies Institute and Centro de Competências de Ciências Exactas e da Engenharia, Universidade da Madeira
Campus da Penteada, 9000-039 Funchal, Madeira, Portugal.
Tel:+351 291-705150/1, Fax: +351 291-705199

Abstract: Artificial Neural Networks have been widely used over three decades. During this period many hardware and software solutions have been developed and today a new user entering the field can make a fast trial to this artificial intelligence solution with commercial software and hardware, instead of developing a solution from scratch and saving a lot of time. This work aims at helping the new and the experienced user even further by sharing the Artificial Neural Network experience in software and hardware, collected through a survey questionnaire about present and past used solutions of software and hardware and prospects for the development of application areas. To further enlighten the reader an analysis with Logistic Regression is performed on the obtained results to extract additional details about the answers obtained from the Artificial Neural Network community.

Keywords: Artificial Neural Networks, System Generator, Neural Network Hardware, Neural Network Software, Linear Regression

1. INTRODUCTION

Artificial Neural Networks (ANN) have been widely used over three decades. In the very beginning researchers were obliged to build their own software to implement the models and eventually build specific hardware for their needs.

Nowadays there are many software and hardware solutions available in the market. Naturally some of these solutions have proved to be more reliable or more fitted to some class of models within the ANN world. To evaluate the use of these solutions, a survey was prepared and submitted to the ANN community.

The survey was presented to the ANN international community through a web site and the invitation was sent through e-mail and the most common mailing lists of the ANN community.

The main objective of the survey questionnaire is to evaluate the use of commercial software and hardware and prospects for the development of application areas among the ANN research community.

For this questionnaire a knowledge discovery approach was taken. A few hypothesis were present at the very beginning, such as to verify if the commercial tools that are well known are in fact more used than any other tool, but the knowledge discovery approach was used to extract the most important conclusion instead of simply verifying previously set hypothesis.

In this paper the results of this survey questionnaire are presented allowing the new and the experienced user to profit from the community experience and select the most tested and proved software and hardware alternatives.

Since a survey questionnaire cannot be easily repeated, it must be carefully prepared. The strategy implemented is depicted in figure 1 through a block diagram.

The first step consists in structuring the survey. It must be done according to the objectives, resources and population. Before presenting the survey, it must be implemented, for example in an informatics platform, and tested.

In the data collection stage it must be ensured that the data is correctly collected and that the platform chosen performs well in collecting and presenting the data. This is closely related to the next block diagram that refers to Access to data. If an informatics platform is used, Data collection and Access to Data are done in the same interface.
Data analysis was done to prepare this paper, using bar and circular graphs and Logistic Regression (LR) to analyse the underlying relations between variables.
This paper is part of the last stage in figure 1.

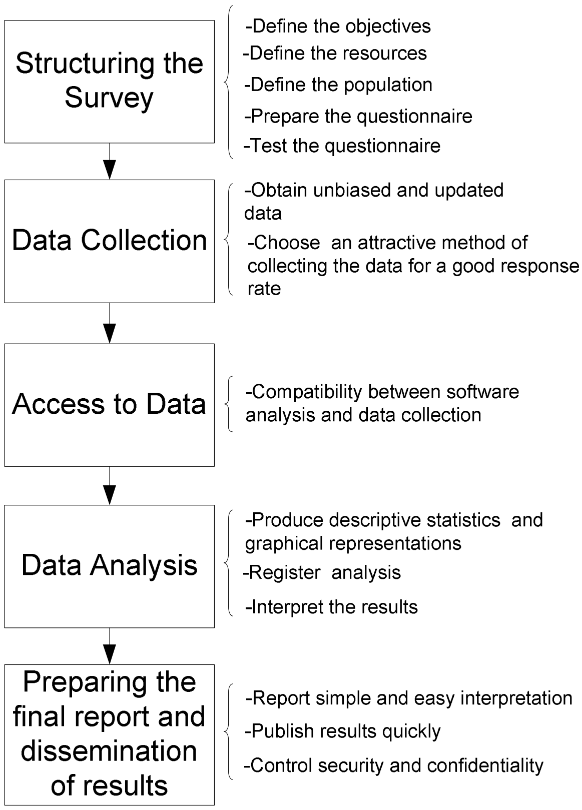

Fig. 1: Block diagram representing the steps taken to implement the survey questionnaire (Alves, N, 2006).

2. SURVEY QUESTIONNAIRE

The survey questionnaire was composed of five parts: Researcher's information, Software information, Application area information, Hardware information and Access to survey results. The last part was only meant to send feedback information to the participants of the survey.
In total the questionnaire was composed of 14 questions.
The questionnaire is available at: https://spreadsheets.google.com/spreadsheet/viewform?formkey=dExPZ1QxUnhLZ2tuVEVacURnTG1jXc6MQ

3. QUESTIONNAIRE ANALYSIS

In a questionnaire, besides reading the direct values of the answers, it is possible to analyse the underlying relations between, in this case, the population and their answers. To do so, in this work, LR (Nemes et. al, 2009) (Peduzzi et. al, 1996) is used. Through LR it is possible to realize which variables explain better the output. Equation 1 shows the logistic function.
The variable z, which is the input of the logistic function, represents the exposure to a set of independent variables, while f(z) represents the probability of a particular outcome, given that set of explanatory variables.

$$f(z) = \frac{e^z}{e^z + 1} \tag{1}$$

The variable z, which is usually defined as shown in equation 2, is a measure of the total contribution of all the independent variables used in the linear model.

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \tag{2}$$

Once the coefficients of the linear regression model are obtained, they describe the size of the contribution of that risk factor. The coefficient is known as odds ratio and can be written as exp(b). The interpretation of exp(b) is as an estimate of the odds ratio between the two related variable. A positive value of the regression coefficient means that the explanatory variable increases the probability of the outcome, while a negative value of the regression coefficient means that the variable decreases the probability of that outcome; a large regression coefficient means that the risk factor has a strong influence in the probability of that outcome, while a near-zero regression coefficient means that the risk factor has little influence on the probability of that outcome (Hasmer and Lemeshow, 1989).
In order to achieve the results, the LR algorithm, at beginning, takes in account all independent variables of equation 2. The LR algorithm is iterative, so at every step it deletes some of the variables that do not have any significant relationship with the output. At the end the LR algorithm leaves only the significant independent variables that have some relationship with z. Sometimes it happens that a z variable does not have any significant relationship with the independent variables in the equation. In that case the conclusion is that there is no relationship between the variable z and the independent variables in the equation. In this study only the significant relationships were presented.
In statistics, a result is called "statistically significant" if it is unlikely to have occurred by chance. More specifically, the significance of a test is the maximum probability of accidentally rejecting a true null hypothesis (a decision known as Type I error - reject the null hypothesis when it is true). For each presented result, the significance value is presented.

4. RESEARCHER'S INFORMATION

The survey sample was composed of 155 elements that accepted the invitation to answer the questionnaire. Figure 2 shows their age distribution. The largest quota of the sample lies in the interval between 41 and 50 years, though globally it is quite well spread.

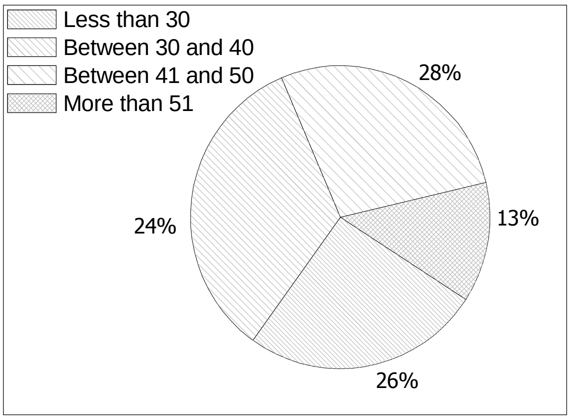

Fig. 2: Age distribution of the survey sample.

Since the sample is composed of researchers it is also relevant to know their academic degree. Figure 3 resumes the highest academic degree of the sample. As can be seen from figure 3, the vast majority of the sample has Doctoral studies and almost 50% of the sample even has post-doctoral studies. Another important question to characterize the sample relates to the research experience.

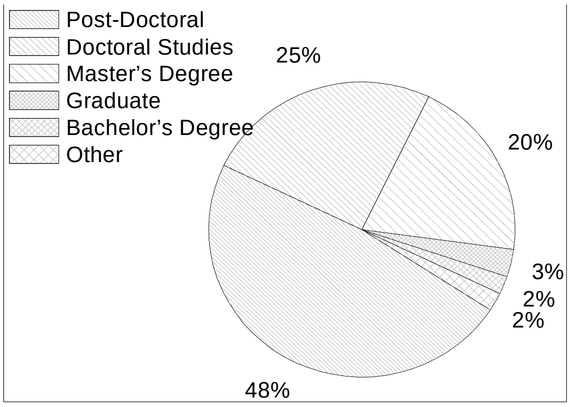

Fig. 3: Highest degree of the survey sample.

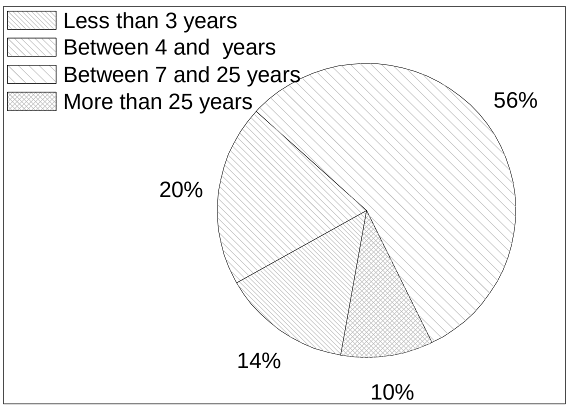

Fig. 4: Number of years involved in research.

Figures 4 and 5 show the number of years of involvement in research in general (figure 4) and in research with ANN in particular (figure 5). One can notice, when comparing figures 4 and 5, that all the intervals diminish except for the less than 3 years interval. This allows concluding that there still new researchers coming to the ANN area.

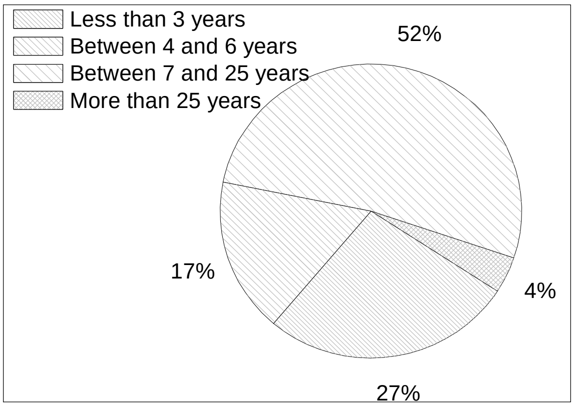

Fig. 5: Number of years involved in research with ANN.

5. SOFTWARE INFORMATION

In this section the focus was placed in knowing the software preferences of the ANN community. In figure 6 it is possible to see the information about the software used more frequently.

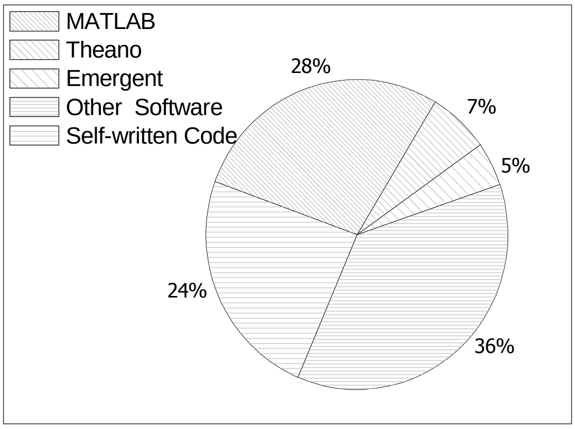

Fig. 6: Software used more frequently among the ANN community.

From figure 6 it is possible to see that Matlab and self-written code take the largest share of choice. There is also a large share of other software being used. This share distributes in the following way:
Neuron and Python (3%), Mathematica, PyBrain, Nest and Brian (2%), Lens, PDP, Weka, NNPred, NeuroSolution, LVQPAK, SOM Toolbox, NNSYSID, Torch, GENESIS, Topographica, Oger, NetLab, AMORE, NEAT, MAPLE and libSVM (1%).
The reasons pointed out for choosing MatLab are: complete software, open-source, powerful and fast processing, easy programming, flexibility and availability of different neural network models. After MatLab, it is possible to find that Theano (7%) and Emergent (5%) are the most used software due to

being simple to use and fast and, only in Theano case, the facility to use the Group Processing Units (GPU). It can also be stated that the use of self-written code is to answer specific research questions, such as study novel types of neural networks which are not available in the commercial software packages.
The LR analysis points out that the researchers that have between 4 and 6 years of experience (in general) do not develop their own software, while the researchers with more than 25 years of experience (in general) choose more to develop their own software (Sig=0.01, Exp(B)=0.26). One side of this question would be to point out that only recently the commercial software reached a development and stability that makes them a valid option. The more experienced researchers had to develop their own software in the first place.
The questionnaire also inquired about the other software used in the community. The answers are presented in figure 7. The answers for this question have a very large value for other software. Among the answer it is possible to find: SNNS, Weka and PLearn (5%), R package and Lens (4%), Brain, Python and NetLab (3%), GENESIS, Brian and SOMPAK (2%), Oger, SPSS, NNtolkit, Nest, Neuron, pybrain, NNSYSID, Neuron, Nest and Joone (1%).

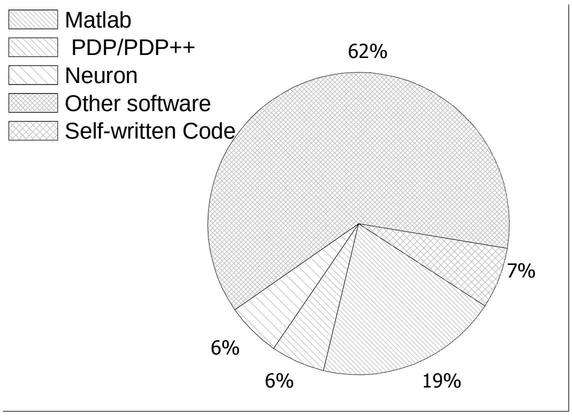

Fig. 7: Other software also used among the ANN community.

From figure 7 is possible to observe that the most used second choice software is Matlab (19%) following of Neuron and PDP/PDP++ (both with 6%).
The LR analysis points out that the researchers of ANN that have an experience between 7 and 25 years are three times more likely to use Matlab than the ANN researchers that have an experience of more than 25 years (Sig=0.068, Exp(B)=2.68). It can also be stretched out that among the PDP users these are more likely to be on the more than 25 years of experience (in general) slot than in the between 7 and 25 years (Sig=0.033, Exp(B)=0.157).
The last question in this section regards the software that best meets the needs. The results are presented in figure 8.
For this last question in this section it is possible to find that the community evaluates Self-written code (34%), Matlab (29%), Emergent (5%) and Brian (3%) as the best solutions.
From the LR analysis it can be pointed out that it is among the researchers of more than 50 years that the development of their own software is more likely to be chosen as the one best meeting the needs (Sig=0.015, Exp(B)=0.216).

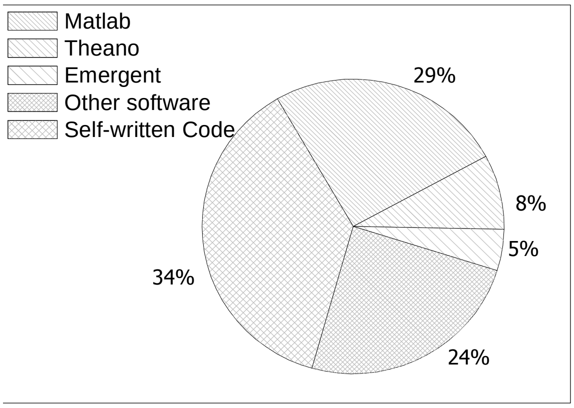

Fig. 8: Software that best meets the needs of the ANN community.

6. APPLICATION AREA INFORMATION

In this section the application areas are evaluated. The first question regards previous developed work. The areas for which the researchers developed work are represented in figure 9. It should be noticed that a multiple selection was possible for this question.

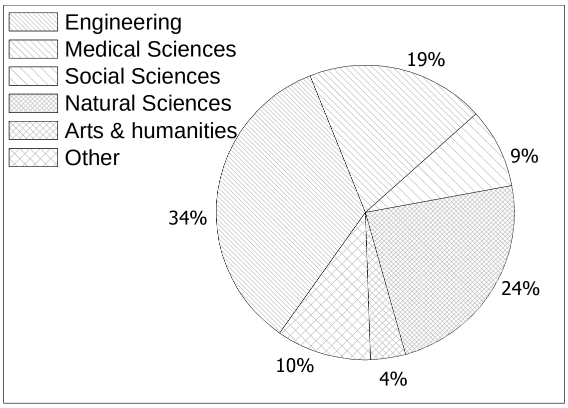

Fig. 9: Areas for previous developed work with ANN.

From this answer is possible to find that the two most selected areas are the Engineering and Natural Sciences areas. It is interesting to note that these areas correspond to the one who developed more ANN, Engineering, and the one that inspired ANN, that is Natural Sciences.
The LR analysis points out that it is the researchers of ANN with an experience between 7 and 25 years that have developed the largest share of the engineering applications and that it is among the researchers with more than 50 years that the largest share of the applications in Social Sciences were developed (Sig=0.036, Exp(B)=2.2).
The next question is about the future use of ANN. The prevision of use of ANN in the future can be seen in figure 10.

Table 1 Resume of the answers regarding specific hardware use for ANN

Hardware for ANN	Yes		No	
	Absolute Frequency	Relative Frequency	Absolute Frequency	Relative Frequency
Question 1 - Need for specific hardware for ANN	61	39%	90	58%
Question 2 – Used Commercial hardware for ANN	32	21%	119	77%
Question 3 – Developed hardware for ANN	15	10%	135	87%
Question 4 – Utility of a toolbox for fast hardware implementation of ANN	88	57%	52	34%

According to the researchers' opinion, the areas that hold the best promise of future development for ANN are: Engineering and Medical Sciences. Natural Sciences will still have an important role.

The LR analysis allows concluding that the share of researchers that elects engineering as the area of bigger growth is larger for the ANN researchers with less than 3 years, when compared with the ANN researchers with more than 25 years of experience (Sig=0.059, Exp(B)=2.821).

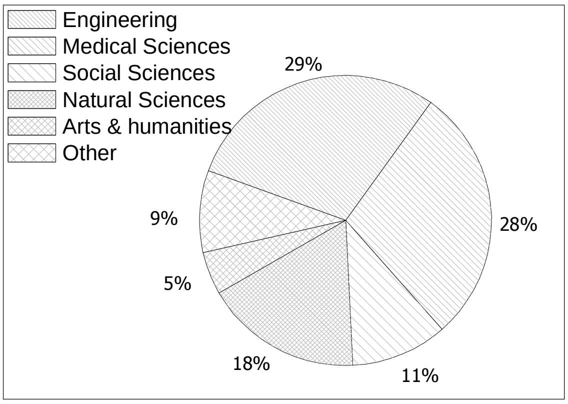

Fig. 10: Areas for future work with ANN.

The same analysis allows concluding that the Natural Sciences area is selected as the one with more potential to grow by the researchers (independently) that have a bachelor's degree (Sig=0.017, Exp(B)=0.042); the researchers (in general) that have an experience between 7 and 25 years (Sig=0.020, Exp(B)=15.98) and the researchers aged between 40 and 50 years (Sig=0.024, Exp(B)=0.096).

7. HARDWARE INFORMATION

For this section the goal was to find about the use of hardware for implementing ANN.

Most of the ANN implementations are done with a Personal Computer (PC). Nevertheless many implementations are implemented with dedicated hardware. The need for leaving the most common implementation of ANN with a PC might arise from a number of reasons: reducing the cost of the implementation, achieving higher processing speed or simpler implementations (Dias et. al, 2004).

To understand the necessity of using specific hardware and the hardware chosen, four questions were placed in the questionnaire:

- Artificial Neural Networks are mostly implemented in Personal Computers. Some applications need specific hardware for their implementation. Have you felt the need for specific hardware?
- Have you used commercial hardware for implementing Artificial Neural Networks?
- Have you developed hardware for Artificial Neural Networks?
- If a toolbox were available for fast hardware implementation of Artificial Neural Networks based on your previously trained network, would you consider using it instead of developing your own hardware?

A resume of the answers obtained is depicted in table 1.

To complement the information, additional details were asked about the commercial hardware used and the development of hardware done by the researcher. The results are presented in figures 11 and 12.

Figure 11 shows very interesting and somehow surprising results. The commercial hardware used by the researchers includes 3 Application Specific Integrated Circuit (ASIC), in a total of 23%, while the remaining 77% of the answers is composed of general purpose hardware: Graphics Processing Units (GPU) and NVIDIA (a graphics processor manufacturer). These solutions are used simply as an hardware accelerator.

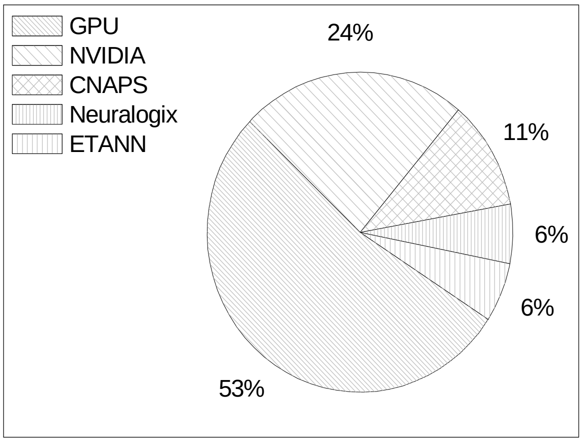

Fig. 11: Commercial hardware used to implement ANN.

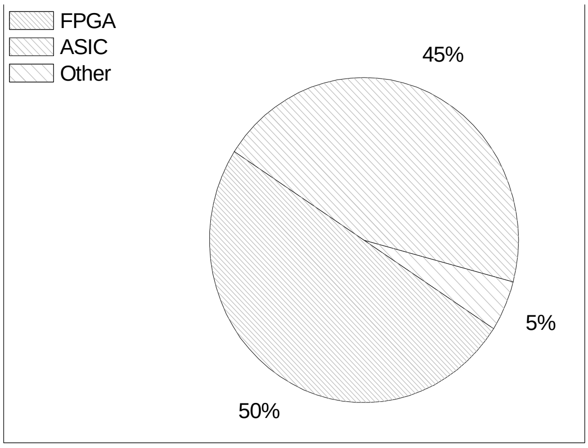

Fig. 12: Platforms used to develop hardware for ANN.

Figure 12 refers to the platforms used by the researchers to develop their hardware. FPGAs are the preferred choice (50%), followed closely by ASICs (45%). The remaining answer refers to an implementation based in Very Large Scale Integration Circuits.

8. CONCLUSION

This paper presents a survey that was submitted to the ANN community. This survey, based in a questionnaire, inquires about software and hardware choices and future prospects for the application areas of ANN.

The results allow extracting important conclusions such as:

-The ANN community is well distributed in the age slots and is still growing with a large share of new researchers and a smaller slot of researchers that are over 51 years;

-More than half of the community has between 7 and 25 years of research both in general and in ANN;

-73% of the sample that answered the questionnaire holds a Ph.D.

-Self-written code and Matlab are the software solutions that best meet the needs of the researchers in ANN and, in the inverse positions, Matlab and Self-written code are the most used software solutions;

-The application area with more implementations developed is engineering and this is also the area that holds the prospect of needing more implementations in the future;

-39% of the ANN researchers felt the need of specific hardware for implementing their networks;

-21% used commercial hardware for ANN;

-10% developed new hardware for ANN;

-The commercial hardware used in ANN is mostly to speed up the calculations;

-FPGAs is the most used platform when developing new hardware solutions, followed closely by ASICs;

-The majority of the researchers feel that a toolbox for fast implementation of ANN in hardware would be usefull.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Portuguese Foundation for Science and Technology for their support for this work through project PEst-OE/EEI/LA0009/2011.

REFERENCES

Dias F. M., Antunes A. and Mota A. M. (2004). Artificial neural networks: a review of commercial hardware, in Engineering Applications of Artificial Intelligence, Vol.17, Issue 8 pp. 945-952.

Nemes S, Jonasson JM, Genell A, Steineck G. 2009 Bias in odds ratios by logistic regression modelling and sample size. BMC Medical Research Methodology 9:56 BioMedCentral

Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996). "A simulation study of the number of events per variable in logistic regression analysis". J Clin Epidemiol 49 (12): 1373–9. PMID 8970487.

Wikipedia, Logistic Regression, available on-line at: http://en.wikipedia.org/wiki/Logistic_regression, consulted in April, 2012.

Alves, N. (2006). Investigação por inquérito, Final report, University of Azores.

Hasmer, D., e Lemeshow, S., Applied Logistic Regression, Wiley, 1989