September 2020
Volume: 17, **No**: 5, pp. 639–649
ISSN: 1741-8984
e-ISSN: 1741-8992
www.migrationletters.com

First Submitted: 1 February 2019 Accepted: 23 May 2020
DOI: https://doi.org/10.33182/ml.v17i5.693

Spending Level of Displaced Population Returned to La Palma, Cundinamarca (2018): A Machine Learning Application

Jenny-Paola Lis-Gutiérrez[1], Mercedes Gaitán-Angulo[2] and Jenny Cubillos-Diaz[3]

Abstract

This research aims to know the variables allowing to predict the spending level of the displacement victims that returned to La Palma, Cundinamarca. For this purpose, a measurement instrument was divided into four sections: characterisation of the population, restitution of economic rights, patterns of economic distribution and, finally, social innovation initiatives. We applied the instrument to 100 participants, and we use different Machine Learning algorithms to know the variables that allow predicting the level of expenses of the displacement victims that returned to La Palma, Cundinamarca. The findings permitted to observe that, at the aggregate level, the Random Forest and the SMV have a prediction capacity higher than 84%.

Keywords: *Displaced population; expenses; victimising events; Colombia armed conflict; machine learning.*

Introduction

Expenditure is a key component in the welfare measure of a population, so, countries use the income or consumption expenses as indicators of poverty. For this reason, many studies have focused on analysing the methods of consumption data collection and the perceived distribution of expenses to obtain levels of poverty in a region (Beegle *et al.*, 2015; Dang & Lanjouw, 2016; Tarozzi, 2007). As a variable, expenditure is more stable than income since it allows a better classification of households and has been considered the standard variable to measure poverty (Deaton, 2005).

Authors such as Shafir (2017) affirm that facts such as poverty, financial challenges, instability of income, expenses and low level of savings occupy a relevant place in populations that have been victims of violence. The context of poverty goes beyond mere survival and is partly a matter of norms and conceptualisation. In this way, as societies advance and standards evolve, things that were once considered a luxury can become common.

Colombia has suffered an internal armed conflict for decades, being the forced displacement one of the most recurrent victimising events in the country (Lagos-Gallego et al., 2017; García-Chavarro, at al., 2018; Lis-Gutiérrez, et al., 2018a, 2018b). The municipality of La Palma, in Cundinamarca region, was, in the year 2015, one of the most affected territories with 13,848 displaced individuals and 2,698 hectares registered as abandoned (Centro Nacional de Memoria Histórica, 2015). However, at present, this municipality is a pioneer in the restitution of lands and return to the territory of the population victim of those crimes.

[1] Jenny-Paola Lis-Gutiérrez, Fundación Universitaria Konrad Lorenz, Colombia. E-mail: jenny.lis@konradlorenz.edu.co.
[2] Mercedes Gaitán-Angulo, Fundación Universitaria Konrad Lorenz, Colombia. E-mail: mercedes.gaitana@konradlorenz.edu.co.
[3] Jenny Cubillos-Diaz, Fundación Universitaria Konrad Lorenz, Colombia. E-mail: jennyk.cubillosd@konradlorenz.edu.co.

The populations that are victims of events such as displacement, and return to their territories, face difficulties in adapting, especially due to labor inclusion issues (Depetris-Chauvin & Santos, 2018; Dabaieh & Alwall, 2018; Contreras, Blaschke, & Hodgson, 2017) and reconciliation and forgiveness (López-López *et al.*, 2019). These changes in their context transformed their social and economic dynamics; so, the monthly low expenses of the displaced population are usually due to informal jobs that accentuate the conditions of vulnerability to which they are exposed (Falla Ramírez, Chávez Plazas, & Molano Beltrán, 2003). This situation does not improve when returning, since populations, in most cases, take several generations to adapt to their territory again (Lis-Gutiérrez *et al.*, 2019). Therefore, the actions framed in the integral reparation of this type of population must be subject to public policies that recognise the complexity of factors that interact in the economic development of their communities (Lacroix & Zufferey, 2019; Johnson *et al.,* 2019).

As a contribution to this subject, the work of Lis-Gutiérrez, et al., (2018c) applies the machine learning principles to measure the incomes of the displaced population in La Palma, finding that the best algorithms were Decision Tree, AdaBoost and Random Forest, with a predictive level between 80% and 90.6%.

Based on these approaches, this research seeks to know the variables that allow predicting the level of expenditures of the population victim of displacement that returned to La Palma, Cundinamarca in 2018. Answering to this question implies knowing the way in which economic dynamics influence social and territorial development, and understanding how the population expenses are related to the historical conditions of violation to their rights.

In this sense, several machine learning or automatic learning algorithms were used, defining machine learning as "the science of giving meaning to data through algorithms and analytical models […] it allows cognitive systems to learn, reason and interact with us in a more natural and personalised way" (IBM, 2018).

The originality of this work is based in two elements: (i) direct work with the community, and (ii) the use of an analysis technique never used so far to understand the relationship between displacement and population expenditures.

Data and Method

In this section, the data, variables and method used for the analysis are presented.

The variables considered emerged from the information obtained in a fieldwork carried out in the municipality of La Palma between July and September 2018, where 100 volunteers were randomly selected, meeting the following conditions: (i) be resident of La Palma, Cundinamarca in 2018, (ii) be part of the displaced population registered in the Unique Registry of Victims (URV), (iii) be over 18 years old, and (iv) belong to socioeconomic levels 1 and 2.

The calculation of the sample was made based on a population size of N = 8,730 inhabitants of the municipality. Having defined the target population range, with the conditions previously predicted, the next step was to calculate the strength of the sample, using the following formula:

$$n = \frac{Z_{\alpha}^{2} Npq}{e^{2}(N-1) + Z_{\alpha}^{2} pq} \quad (1)$$

Where

N: is the size of the population or universe (total number of potential respondents) = 8.730

$Z\alpha$ = 95%

e = 10%, is the desired sample error.

p = 50% proportion of individuals who meet the study conditions in the population, given that the exact parameter is unknown, 50% is used.

q = 50% proportion of individuals who do not meet the study conditions in the population.

n: sample size

Only 95 individuals were required to guarantee an error margin of 10% and confidence level of 95%. Given the sample of 100 individuals, the confidence level is 95%, and the error margin is reduced to 9.75%. The instrument applied (survey) received a content validation from three expert judges, and a pilot test was applied to 12 people for testing the understanding level of the instrument. The survey used to collect the information was divided into three sections. The first one was oriented to know the socio-demographic characteristics of the population (age, level of education, place of birth, etc.). The second section was related to questions about the rights restitution process to communities that suffered displacement. And the last one sought to know the type of monthly economic expenses and the economic conditions of the household.

The instrument was applied by two volunteers with experience in handling people in the field. The survey was conducted with presence of the interviewer who completed the instrument, during a period of 23,5 minutes. Each of the participants carried out the survey following the following sequence: (i) reading of the informed consent was completed; (ii) application of the filter for validation of inclusion and exclusion criteria; (iii) reading the general instructions; and (iv) application of the instrument.

The instrument was applied as established in Law 1090 of 2006 (Congreso de la República de Colombia, 2011) about the professional practice of the psychologist, from the ethical and procedural point of view, meeting the following criteria: (i) provide clear information about the type of research carried out, the people responsible for it and the institution for which it was being carried out; (ii) inform about the confidential nature of the information provided and the exclusive use of the investigators; (iii) provide sufficient information about the purpose of the investigation; (iv) inform the average duration of the test, the tasks to be performed, and the non-existence of risk of secondary effects or health damages to the participants; (v) explain the manifest of voluntary participation in the investigation; and (vi) inform that the test could be abandoned at any time on a voluntary basis.

The variables that were analysed in the document were the following (Table 1).

Table 1. Instrument variables

Variable	Typology
What is your place of birth?	Categorical
According to the utility bills of your home, what is your stratum? (This information is found on the service invoice)	Categorical
How old are you?	Numeric
What sex do you belong to?	Categorical

Table 1. Continued.

Actually lives in	Categorical
What is your occupation?	Categorical
How many people make up your home?	Numeric
Of these people, how many are financially dependent on you?	Numeric
What is the maximum level of education you have reached?	Categorical
In what year was you admitted as a victim at Unique Registry of Victims (URV)	Numeric
Why is it registered with the Unique Registry of Victims (URV)?	Categorical
In what year did you have to leave the municipality?	Date
To date, have you received any comprehensive reparation measure?	Categorical
If your answer was *yes*, which of the following measures have been part of this comprehensive repair? [financial compensation, repair individual or collective]	Categorical
If the answer was *financial compensation,* could you specify *yes*?	Categorical
If your answer was "You have already used these resources in other investments" could you describe in which investments?	Categorical
Have you received any kind of accompaniment from the state in the investment or expenditure of these resources?	Categorical
If your answer was *yes*, please describe the type of accompaniment	Categorical
In what year were you beneficiary of the repair?	Date
Was this repair Individual or Collective?	Categorical
Normally, how much are the monthly expenses of this home? [COP 1=0-500.000; 2=501.000 -1.000.000; 3=1.001.000-2.000.000; 4= >2.001.000]	Categorical
Normally, how much do the monthly expenses of the household correspond to? [COP 1=0-500.000; 2=501.000 -1.000.000; 3=1.001.000-2.000.000; 4= >2.001.000]	Categorical
Does any member of the household have credits or debts with entities, relatives, friends or people?	Categorical
If your answer to the previous question was *Yes*, specify the value of the credit or debt	Numeric
Have you used the economic resources of credit in a business initiative?	Categorical

Source: own elaboration.

For information processing, several methods (algorithms) of supervised learning were applied, which are summarised in Table 2.

Table 2. Supervised learning algorithms

Method	Synthesis
AdaBoost	It is an adaptive learning algorithm that combines weak classifiers and adapts to the hardness of each training sample, achieving robust classifiers. Allows binary and multi-class classification.
Random Forest	Uses a set of decision trees for the classification and projection. It is a non-parametric procedure which can be used when: (i) there are correlated variables; (ii) few data; (iii) complex interactions; (iv) missing data.
Support Vector Machines (SVM)	Allows the binary classification and multi-class. Makes use of regression analysis and classification.
Neural network	Makes use of the algorithm for multi-layer perceptron (MLP) with retropropagation. Its advantage is that nonlinear models can be learned.
KNN	This mechanism is for the recognition of non-parametric patterns; it is based on the nearest training instances, being an algorithm of k nearest neighbors. It is also known as lazy learning.

Table 2. Continued.

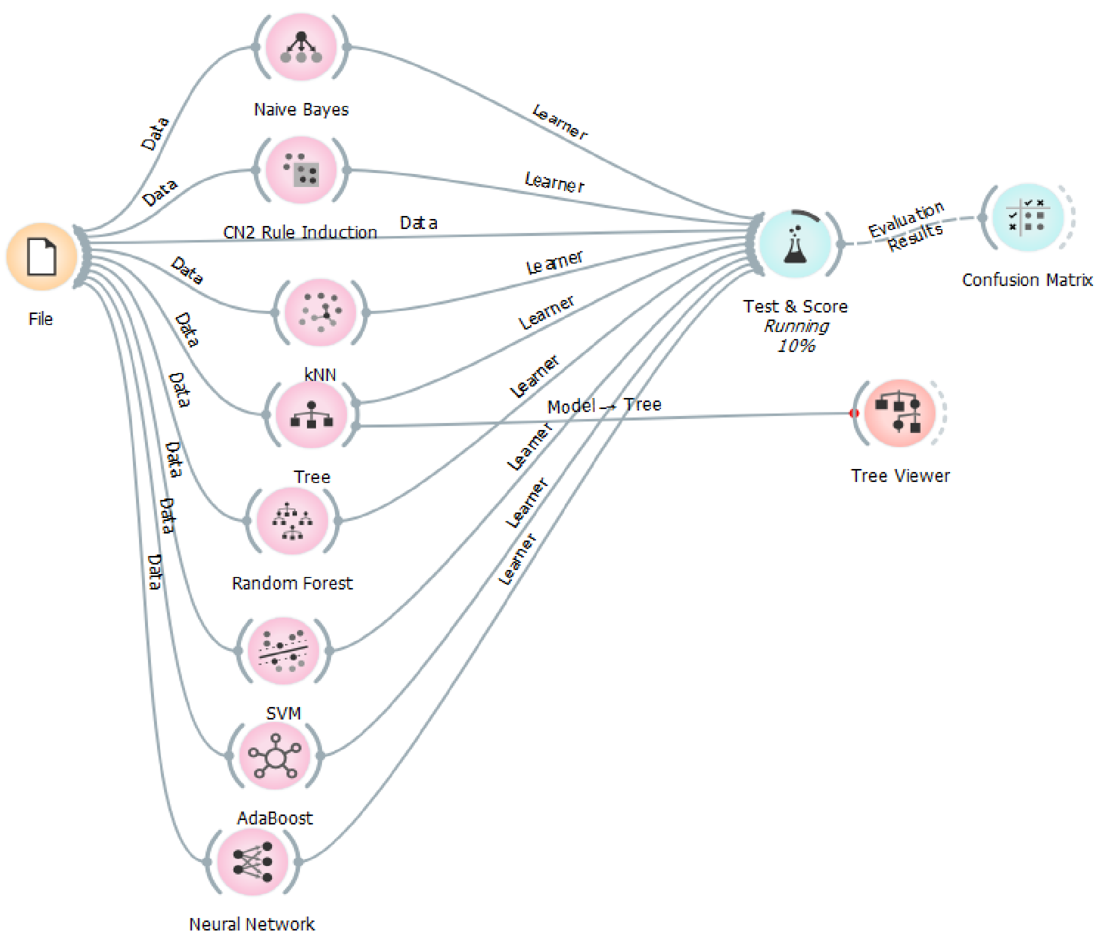
Naive Bayes	A quick and simple probabilistic classifier based on the Bayes theorem, with the assumption of independence of features. It is assumed that the absence or presence of a characteristic is not related to the presence or absence of another property. Its advantage is that it requires little data for training.
Learning algorithm of the decision tree	It is used for discrete and continuous data. It is based on the prediction of the value of a target variable based on various input variables.

Source: Lis-Gutiérrez & Aguilera-Hernández (2019) and Moros et al., (2019).

Results

In Figure 1, the different learning and prediction algorithms used are shown, and in Figure 2, the precision calculations of the same. It should be noted that the use of the random cross validation option was used. This method consisted of randomly dividing the training and test data set, by means of 10 iterations. The adjustment is obtained from the arithmetic mean of the values obtained for each of the iterations. Figure 3 shows the confusion matrix for each algorithm.

Figure 1. Image of the model representation using Orange



Source: own elaboration using Orange (Demsar, et al., 2013).

Figure 2. Calculation of learning and prediction algorithms for the complete sample.

Test & Score

Settings

Sampling type: Stratified 20-fold Cross validation
Target class: Average over classes

Scores

Method	AUC	CA	F1	Precision	Recall
Random Forest	0.829	0.870	0.863	0.872	0.870
SVM	0.842	0.790	0.765	0.844	0.790
AdaBoost	0.779	0.790	0.792	0.796	0.790
Tree	0.753	0.800	0.795	0.792	0.800
Naive Bayes	0.817	0.740	0.737	0.750	0.740
Neural Network	0.760	0.740	0.732	0.730	0.740
CN2 rule inducer	0.689	0.710	0.701	0.700	0.710
kNN	0.594	0.600	0.495	0.499	0.600

Source: own elaboration using Orange (Demsar, et al., 2013).

Figure 3. Confusion matrix for each algorithm (complete sample).

Data

Data instances: 100
Features: ID, ¿Cuál es su lugar de nacimiento?, ¿Según las facturas de servicios públicos de su casa que estrato es?, ¿Cuántos años tiene?, ¿A qué sexo pertenece?, Actualmente vive en, ¿Cuál es su ocupación?, ¿Cuántas personas conforman su hogar?, ¿De estas personas cuantas dependen económicamente de usted?, ¿Cuál es el nivel máximo de escolaridad titulada que usted alcanzó?, ¿En qué año fue admitido como víctima ante el RUV?, ¿Por qué hecho se encuentra registrado ante el RUV?, ¿En qué año tuvo que desplazarse del municipio?, ¿A la fecha cuenta con alguna medida de reparación integral?, ¿Si su respuesta fue si ¿Cuál de las siguientes medidas han hecho parte de esta reparación integral?, ¿Si la respuesta fue indemnización económica podría especificar si?, Si su respuesta fue "Ya ha hecho uso de estos recursos en otras inversiones" podría describirnos en cuáles inversiones, ¿Ha recibido alguna clase de acompañamiento por parte del estado en la inversión o el gasto de estos recursos?, ... (total: 41 features)

Confusion Matrix

Confusion matrix for Naive Bayes (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	84.2 %	20.0 %	39.1 %	**61**
	1,000,001 a 2,000,000	0.0 %	60.0 %	0.0 %	**12**
	500,001 a 1,000,000	15.8 %	20.0 %	60.9 %	**27**
	∑	**57**	**20**	**23**	**100**

Figure 3. Continued.

Confusion Matrix

Confusion matrix for CN2 rule inducer (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	74.6 %	23.1 %	40.0 %	**61**
	1,000,001 a 2,000,000	4.5 %	69.2 %	0.0 %	**12**
	500,001 a 1,000,000	20.9 %	7.7 %	60.0 %	**27**
	∑	**67**	**13**	**20**	**100**

Confusion Matrix

Confusion matrix for kNN (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	63.0 %	20.0 %	66.7 %	**61**
	1,000,001 a 2,000,000	12.0 %	20.0 %	0.0 %	**12**
	500,001 a 1,000,000	25.0 %	60.0 %	33.3 %	**27**
	∑	**92**	**5**	**3**	**100**

Confusion Matrix

Confusion matrix for Tree (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	84.6 %	0.0 %	26.1 %	**61**
	1,000,001 a 2,000,000	0.0 %	83.3 %	8.7 %	**12**
	500,001 a 1,000,000	15.4 %	16.7 %	65.2 %	**27**
	∑	**65**	**12**	**23**	**100**

Figure 3. Continued.

Confusion Matrix

Confusion matrix for Random Forest (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	85.7 %	0.0 %	5.0 %	**61**
	1,000,001 a 2,000,000	0.0 %	100.0 %	10.0 %	**12**
	500,001 a 1,000,000	14.3 %	0.0 %	85.0 %	**27**
	∑	**70**	**10**	**20**	**100**

Confusion Matrix

Confusion matrix for SVM (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	74.4 %	0.0 %	0.0 %	**61**
	1,000,001 a 2,000,000	6.1 %	100.0 %	0.0 %	**12**
	500,001 a 1,000,000	19.5 %	0.0 %	100.0 %	**27**
	∑	**82**	**7**	**11**	**100**

Confusion Matrix

Confusion matrix for AdaBoost (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	83.9 %	0.0 %	32.1 %	**61**
	1,000,001 a 2,000,000	0.0 %	100.0 %	7.1 %	**12**
	500,001 a 1,000,000	16.1 %	0.0 %	60.7 %	**27**
	∑	**62**	**10**	**28**	**100**

Figure 3. Continued.

Confusion Matrix

Confusion matrix for Neural Network (showing proportion of predicted)

		Predicted			
		0 a 500,000	**1,000,001 a 2,000,000**	**500,001 a 1,000,000**	**∑**
Actual	**0 a 500,000**	77.6 %	0.0 %	40.9 %	**61**
	1,000,001 a 2,000,000	1.5 %	90.9 %	4.5 %	**12**
	500,001 a 1,000,000	20.9 %	9.1 %	54.5 %	**27**
	∑	**67**	**11**	**22**	**100**

Source: own elaboration using Orange (Demsar, et al., 2013).

After applying 8 supervised learning algorithms to the survey information, it was possible to identify that:

When considering the average results, the following algorithms have a predictive capacity between 84.4% and 87.2%: Random Forest and SVM; and between 75% and 79.6% had a predictive capacity AdaBoost, Tree, Naive Bayes (Figure 2).

The algorithm with the least predictive capacity was the KNN.

Considering the Figure 3, it is possible to indicate that:

The best algorithms for the prediction of the range of costs between 0 and 500,000 pesos were: (i) Random Forest: 85.7%; (ii) the decision tree: 84.6%; (iii) Naive Bayes: 84.2%; (iv) AdaBoost: 83.9%.

The best algorithms for the prediction of range of expenditure between 500,001 and 1,000,000 pesos were: (i) SVM: 100%; (ii) Random Forest: 85%.

The best algorithms for the prediction of range of expenditure between 1,000,001 and 2,000,000 pesos were: (i) Random Forest: 100%; (ii) SVM: 100%; (iii) AdaBoost: 100%; (iv) Neural Networks: 90.9%; (v) decision tree: 83.3%

The others algorithms do not predict any of the three ranges with an accuracy higher than 80%, therefore, should not be taken into account: KNN y CR2.

Now, the application of the decision tree displaying the variables that allow to predict in 84.6% the population that gets between 0 and 500,000 monthly. Its characteristics are the following: (i) the level of education ranges between primary, secondary or technical support; (ii) have received compensation by the State; (iii) have generally used the resources for housing arrangements and purchase of animals for the farm. Likewise the application of the decision tree displaying the variables that allow to predict in 83.3% the population that gets between 1,000,001 and 2,000,000 monthly. The level of income ranges between 500,000 and 2,000,000 pesos, but most of those who earn between 500,001 and 1,000.000 are located in a rural area.

Conclusions

The variables that allowed explaining the behaviour of the expenses of the victims who returned to La Palma, Cundinamarca were identified. It was possible to establish that the three best supervised learning algorithms, for the average, were at the aggregate level the Random Forest and the SMV emblem with a capacity of prediction, higher than 84%. As in the work of Lis-Gutierrez et al, (2018c), the algorithm with less capacity of prediction was the KNN. The Random Forest, the decision tree, Naive Bayes and AdaBoost: predict the initial income range with a level higher than 83.9% (between 0 and 500,000 pesos, i.e., between 0 and 167 dollars). For the prediction of middle income range (between 500,001 and 1,000.000, i.e. between 167 and 334 monthly dollars), the SVM had a predictive capacity of 100%, followed by Random Forest. With regard to the prediction of middle income range (between 1,000,001 and 2,000.000, i.e. between 334 and 668 monthly dollars), the SVM, the Random Forest and AdaBoost, had a predictive capacity of 100%, followed by the neural networks and the decision tree.

Among the recommendations for future studies are: (i) enlarge the sample, in order to have more predictive power and reduce the error margin from 9.75% to 5%; (ii) replicate this study in other populations, and check if the algorithms identified efficiently predict the income levels of the returning population. In either case, communicating of the findings to the community is required.

Acknowledgements

The authors wish to express their gratitude to the community of La Palma (Cundinamarca).

References

Beegle, K., De Weerdt, J., Friedman, J., & Gibson, J. (2015). Methods of household consumption measurement through surveys: Experimental results from Tanzania. Washington: Banco Mundial.

Centro Nacional de Memoria Histórica. (2015). *Una nación desplazada: informe nacional del desplazamiento forzado en Colombia*. Bogotá: CNMH–UARIV. Retrieved from: http://www.centrodememoriahistorica.gov.co/descargas/informes2015/nacion-desplazada/una-nacion-desplazada.pdf.

Congreso de la República de Colombia (2011). *Ley 1090 de 2006*. Diario Oficial 46.383. Bogotá: Imprenta Nacional.

Contreras, D., Blaschke, T., & Hodgson, M. E. (2017). Lack of spatial resilience in a recovery process: Case L'Aquila, Italy. *Technological Forecasting and Social Change*, 121, 76-88.

Dang, H. A. H., & Lanjouw, P. F. (2016). *Toward a new definition of shared prosperity: A dynamic perspective from three countries.* In Inequality and growth: Patterns and policy (pp. 151-171). Londres: Palgrave Macmillan.

Dabaieh, M., & Alwall, J. (2018). Building now and building back. Refugees at the centre of an occupant driven design and construction process. *Sustainable cities and society*, 37, 619-627.

Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in a poor world). *Review of Economics and statistics*, 87(1), 1-19.

Demsar, J., Curk, T., Erjavec, A., Gorup C, Hocevar, T., Milutinovic, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M., Zupan, B. (2013). Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14(Aug), 2349−2353.

Depetris-Chauvin, E. y Santos, R. (2018). Unexpected guests: The impact of internal displacement inflows on rental prices in Colombian host cities. *Journal of Development Economics,* 134, 289-309.

Falla Ramírez, U., Chávez Plazas, Y. A., & Molano Beltrán, G. (2003). Desplazamiento forzado en Colombia. Análisis documental e informe de investigación en la Unidad de Atención al Desplazado (UAID). *Tabula Rasa, 1*, 221-234. Retrieved from: http://www.redalyc.org/pdf/396/39600111.pdf

García-Chavarro, A., Lis-Gutiérrez, J-P., Rincón-Vásquez, J.C., Gaitán-Angulo, M., Cubillos-Diaz, J., Mojica-Sánchez, L. (2018). Caracterización de los hechos victimizantes por región en las víctimas del conflicto armado colombiano.

En J.P. Lis-Gutiérrez, M. Gaitán-Angulo, J. Cubillos-Díaz, L. Mojica Sánchez y L.E. Malagón. *Conflicto y construcción de paz*. Villavicencio: Unimeta.

IBM (2018). Machine learning. Retrieved from: https://www.ibm.com/analytics/es/es/technology/machine-learning/

Johnson, T., von Meding, J., Gajendran, T., & Forino, G. (2019). Disaster vulnerability of displaced people in Rakhine state, Myanmar. In Resettlement Challenges for Displaced Populations and Refugees (pp. 81-91). Springer, Cham.

Lacroix, J., & Zufferey, J. (2019). A Life Course Approach to Immigrants' Relocation: Linking Long-and Short-distance Mobility Sequences. Migration Letters, 16(2), 283-300.

Lagos-Gallego, M., Gutierrez-Segura, J. C., Lagos-Grisales, G. J., & Rodriguez-Morales, A. J. (2017). Post-traumatic stress disorder in internally displaced people of Colombia: An ecological study. *Travel medicine and infectious disease, 16*, 41-45.

Lis-Gutiérrez, J. P., Rincón-Vásquez, J.C., Gaitán-Angulo, M., Cubillos Diaz, J., Vargas, C. (2019). *Hechos victimizantes en Colombia: antes, durante y después de la firma del acuerdo de paz en La Habana*. En: G.A. Campos Avendaño, M.A. Castaño Hernández, M. Gaitán Ángulo y V. Sánchez Mendoza (Compiladores). Diálogos sobre investigación: avances científicos Konrad Lorenz (25-50). Bogotá: Konrad Lorenz Editores. Disponible en: https://blogs.konradlorenz.edu.co/files/book_interactivo_dialogos_en_investigacion.pdf

Lis-Gutiérrez, J.P., Cubillos Díaz, J., Rincón, J.C., Gaitán-Angulo, Lis-Gutiérrez, M., Henao, C., y Balaguera, M.I. (2018a). Ingreso de la población víctimas de desplazamiento que retornó a la Palma Cundinamarca (2018): una aplicación de machine learning. En: J.P. Lis-Gutiérrez, C. Henao y L.E. Malagón-Castro. Técnicas de análisis cuantitativo aplicadas a las ciencias contables y económicas (240-260). Villavicencio: Unimeta.

Lis-Gutiérrez, J-P., Rincón-Vásquez, J.C., Cubillos-Diaz, J., Gaitán-Angulo, M., Mojica-Sánchez L. (2018b). Caracterización de los hechos victimizantes por género en las víctimas del conflicto armado Colombiano. En: Lis-Gutiérrez, et al., *Conflicto y Construcción de paz.* Villavicencio: Unimeta.

Lis-Gutiérrez, J.P. y Aguilera-Hernández, D. (2019). Análisis departamental de las demandas al Ejército colombiano por sus integrantes en calidad de víctimas. En: G. Barbosa Castillo, M. Correa, y A. Ciro Gómez (eds.), Análisis de las demandas de los integrantes del Ejército en calidad de víctimas: una aplicación de "machine learning" (pp. 324-352). Bogotá: Universidad Externado de Colombia.

López López, W., Correa-Chica, A., Sierra-Puentes, M.C., Castañeda Polanco, J.C., Fernández Miranda, G., Durán Jaramillo, M. & Castro-Abril, P. (2019 In press). Armed Conflict Conceptualization of Forgivenness, ReConciliation and Peace: Children's Voices in the Armed Conflict Context. In N. Balvin, & D. J. Christie (Eds.) Children and Peace: From Research to Action. New York: Springer: Peace and Psychology Book Series.

Moros Ochoa, M.A., Lis-Gutiérrez, J. P., Castro Nieto, G.Y., Vargas, C. y Rincón-Vásquez, J.C. (2019). La percepción de calidad de servicio como determinante de la recomendación: una predicción mediante inteligencia artificial para los hoteles en Cartagena. En: G.A. Campos Avendaño, M.A. Castaño Hernández, M. Gaitán Ángulo y V. Sánchez Mendoza (Compiladores). Diálogos sobre investigación: avances científicos Konrad Lorenz (pp. 143-162). Bogotá: Konrad Lorenz Editores. Disponible en https://blogs.konradlorenz.edu.co/files/book_interactivo_dialogos_en_investigacion.pdf

Sierra-Puentes, M., & Correa-Chica, A. (2019 In press). Aspectos socioeconómicos de la población desplazada en Colombia: revisión sistemática. Suma Psicológica, 26(2) http://dx.doi.org/10.14349/sumapsi.2019.v26.n2.1

Shafir, E. (2017). Decisions in poverty contexts. *Current opinion in psychology*, 18, 131-136.

Tarozzi, A. (2007). Calculating comparable statistics from incomparable surveys, with an application to poverty in India. *Journal of business & economic statistics*, 25(3), 314-336.