Universidad Continental

FACULTAD DE INGENIERÍA

Escuela Académico Profesional de Ingeniería Mecatrónica

Tesis

An Algorithm to Estimate the Risk of Child Labor

Ruben Aldair Andamayo Alcantara
Ricky Bryan Quiñones Fabian
Abel Jesus Inga Lopez
Jaime Antonio Huaytalla Pariona
Jimmy Alberth Deza Quispe

Para optar el Título Profesional de
Ingeniero Mecatrónico

Huancayo, 2022

Universidad Continental

INFORME DE CONFORMIDAD DE ORIGINALIDAD DE TESIS: EN FORMATO ARTÍCULO CIENTÍFICO

A	:	Felipe Gutarra Meza
		Decano de la Facultad de Ingeniería
DE	:	Jaime Antonio Huaytalla Pariona
		Asesor de tesis
ASUNTO	:	Remito resultado de evaluación de originalidad de tesis
FECHA	:	20 de Febrero del 2024

Con sumo agrado me dirijo a vuestro despacho para saludarlo y en vista de haber sido designado asesor de la tesis titulada: "An algorithm to estimate the risk of child labor", perteneciente al/la/los/las estudiante(s) Ricky Bryan Quiñones Fabian, Ruben Aldair Andamayo Alcantara y Abel Jesus Inga Lopez, de la E.A.P. de Ingeniería Mecatrónica; se procedió con la carga del documento a la plataforma "Turnitin" y se realizó la verificación completa de las coincidencias resaltadas por el software dando por resultado 0 % de similitud (informe adjunto) sin encontrarse hallazgos relacionados a plagio. Se utilizaron los siguientes filtros:

- Filtro de exclusión de bibliografía   SI [X]   NO [ ]

- Filtro de exclusión de grupos de palabras menores   SI [X]   NO [ ]
  (N° de palabras excluidas: 7)

- Exclusión de fuente por trabajo anterior del mismo estudiante   SI [X]   NO [ ]

En consecuencia, se determina que la tesis en formato artículo científico constituye un documento original al presentar similitud de otros autores (citas) por debajo del porcentaje establecido por la Universidad.

Recae toda responsabilidad del contenido la tesis en formato artículo científico sobre el autor y asesor, en concordancia a los principios de legalidad, presunción de veracidad y simplicidad, expresados en el Reglamento del Registro Nacional de Trabajos de Investigación para optar grados académicos y títulos profesionales – RENATI y en la Directiva 003-2016-R/UC.

Esperando la atención a la presente, me despido sin otro particular y sea propicia la ocasión para renovar las muestras de mi especial consideración.

Atentamente,

DECLARACIÓN JURADA DE AUTORÍA

El presente documento tiene por finalidad declarar adecuada y explícitamente el aporte de cada estudiante en la elaboración del artículo de investigación a ser utilizado para la sustentación de tesis: formato de artículo científico.

Yo: Ricky Bryan Quiñones Fabian, con Documento Nacional de Identidad N° 71700321; teléfono 969990241; estudiante de la Escuela Académico Profesional de Ingeniería Mecatrónica.

Yo: Ruben Aldair Andamayo Alcantara, con Documento Nacional de Identidad N° 71431387; teléfono 969192633; estudiante de la Escuela Académico Profesional de Ingeniería Mecatrónica.

Yo: Abel Jesus Inga Lopez, con Documento Nacional de Identidad N° 60354789; teléfono 995600508; estudiante de la Escuela Académico Profesional de Ingeniería Mecatrónica.

Yo: Jaime Antonio Huaytalla Pariona , con Documento Nacional de Identidad N° 48560801; teléfono 990223332; docente de la Escuela Académico Profesional de Ingeniería

Yo: Jimmy Alberth Deza Quispe , con Documento Nacional de Identidad N° 76845782; teléfono 980655854; docente de la Escuela Académico Profesional de Ciencias de la Empresa

Ante Usted, con el debido respeto me presento y expongo:

Declaramos que hemos participado en la ideación del problema, recolección de datos, elaboración y aprobación final del artículo científico.

An algorithm to estimate the risk of child labor

Decision Science Letters

An algorithm to estimate the risk of child labor

Ricky Bryan Quiñones Fabian[a*], Ruben Aldair Andamayo Alcantara[a], Abel Jesus Inga Lopez[a], Jaime Antonio Huaytalla Pariona[a] and Jimmy Alberth Deza Quispe[a]

[a]Universidad Continental, Peru

CHRONICLE	ABSTRACT
	In developing countries, child labor has become a significant problem with adverse effects in the present and future for society and individuals. There are many causes that obligate children to abandon school and start working. Economic, social, familiar, and personal problems can expel children from school, inhibiting them from living appropriately. Polls like the ENAHO in Peru tried to recollect relevant data as much as possible to explain this problem. With many variables, it is necessary to have a methodology to build an algorithm with enough explanatory power to explain the situation. Therefore, this research elaborated an algorithm through Lasso to proportionate a statistical explanation of child labor. Due to the type of data, the regression was logistic.

1. Introduction

Child labor is a global issue with adverse effects on youngster development (Morales & Vargas, 2019). Besides, exploitation in many forms is a constant threat (Bureau of International Labor Affairs, 2019). Consequently, people and organizations are making efforts to eradicate child labor. However, today children must work to survive (Velazco, 2016). Hence, in Latin America, 7.3% of children and adolescents work (Ruiz, 2019). In Peru, the Andean region concentrates 65.3% of children labor cases (Instituto Nacional de Estadística e Informática, 2011). Poverty is the main reason for child labor and probably school dropout. For instance, Huebler (2008) found an inverse relationship between child labor and school attendance, where poverty was an important motivation. Additionally, Adam et al. (2016) encountered that poverty pushed children to generate income immediately to survive. However, income level is not the only factor for school dropout. For instance, Abud (2015) found that family culture was essential to understanding the value of education as a key to better opportunities in the future. Consequently, there are economic, social, spatial, and temporal grounds (Revenberg, 2015).

There are negative psychological and physical consequences of child labor (Uddin et al., 2011). Those effects are long-lasting for the children. Moreover, a country cannot accumulate human capital due to child labor. Working hours force children to start missing classes and eventually drop out of school (Khan et al., 2011). Hence, the postponement of skills acquisition in years of education damages the country's economy and cuts its potential because of the lack of knowledge and technology (Edmonds, 2015)). Therefore, it is necessary to identify and understand the factors that push children to start working and leaving school to elaborate an algorithm for prevention strategies.

2. Literature review

Academic research has studied the determinants of school dropout and children's work. For instance, Morales and Vargas (2019) explored this phenomenon in Bolivia. Hence, they employed a probit model with family and social factors. Therefore, the study uncovered that family income, education degree, structure, and Bolivia's regional differences, were significant for children to continue school.

Alike, Hernández and Vargas (2016) harnessed a Mexican survey to explore the association between children's employment and school dropout in urban areas. The researchers found that longer than 20 hours a week working hours and low salaries pushed students to abandon school. Furthermore, children's family characteristics and educational performance influenced the probability of school dropout.

Moreover, Salata (2019) employed the multivariate logit model to investigate the causes of school dropout in Brazil. It showed that long working hours significantly increase the chances of dropping out of school. Moreover, the analysis found that socioeconomic status influenced school attendance. Then, the children's family cultural characteristics influenced their decision to stay at school.

Machado et al. (2015) also studied the same phenomenon in Brazil in a specific region. That research employed regression and Logit models to obtain its results. The study found that male students leave school because of poor academic performance and low motivation, while pregnancy was the main reason for women. Besides, the research found that students had low motivation because of economic difficulties. Hence, they prefer to contribute to their family's current necessities by working instead of getting an education.

Asian countries also deal with absenteeism and child labor. For instance, Khan et al. (2011) analyzed the factors behind child labor in rural Bangladesh. They utilized the Bengali household survey along with bivariate models. Hence, the research encountered inverse relationships between parents' academic level, family income, and gender with child labor. Furthermore, they located a higher mother study level effect than the father one regarding youngsters dropout. Moreover, in impoverished families, longer work hours make study hours smaller. Furthermore, the study found that girls are more eager to work longer hours than men, vaccinating more absenteeism and increasing their probability to drop out of school.

Previous studies employed national household surveys as their data source. Also, they encountered that family income and educational level, along with child gender, were crucial factors of school dropout and child labor. However, a complete analysis should also consider the environmental characteristics where the children grow up. For instance, Alcázar (2009) found that rural areas had a higher school dropout ratio than urban ones.

In Peru, Pariguana (2011)explored the determinants of school dropout and child labor. Hence, the research employed the bivariate probit model with the 2007 child labor survey. The analysis used the children's family expenditure, education level, head gender, language, living area, and children under care. Moreover, they considered students' age and gender. The research discovered that family earnings and living zones were determinants. The study found that many rural families had low incomes due to their head members' lack of education. Furthermore, many of them cannot even speak or write Spanish. Hence, rural students decide to work instead of studying to survive in a hostile environment.

Consequently, this research will investigate school dropout and labor causes between five and fourteen-year-old students. Hence, this research will build an algorithm to examine those factors with further validation through the Logit analysis. It is necessary to add that this research will study the phenomenon in the central region of Peru since they are the epicenter of school dropouts (Miranda, 2018).

3. Theoretical background

Skoufias's model highlights the benefits of investing in human capital and education for the future (Skoufias, 2005). This model proposes that the child's future income will depend on his accumulated knowledge and skills. In the case of child labor, when the child works, he only receives a fraction of the income that the same person would receive if they accumulated human capital. Therefore, it is not easy to operate and study simultaneously, so in the case of children, they are sacrificing future income (Pariguana, 2011). Likewise, some factors can impact the decision to study or work for children (Pariguana, 2011). Based on the literature, these factors can be household expenditure, home chief gender, family educational level, children's gender, age, and repetition. Therefore, the current study will evaluate the factors that can affect the decision of school dropouts to work or keep studying. The following section will provide the theoretical scope.

3.1 School dropout and child labor

According to Skoufias's model, education forms human capital, as stated before. The economic literature defined human capital in many ways. For instance, Becker & Tomes (1986) described human capital as the set of capabilities that an individual acquires by accumulating general or specific knowledge.

A child is attending school for practical purposes when absences do not exceed three months, regardless of academic performance. Otherwise, it is going to be considered a situation of school dropout. This concept applies to the essential research database of the ENAHO of 2017. The current research will assume that working and studying are mutually exclusive.

3.2 Child work

Child labor interpretation varies according to the country's legal and social norms. Therefore, Velazco (2016) defines child labor as any activity that inhibits children from living and enjoying their potential and dignity, harmful to their development. Therefore, child labor is always an activity that threatens the health and personal development (Velazco, 2016). Otherwise, those activities can be positive for the children's development. For instance, these activities can include helping parents with household chores, support in a family business, or tasks done outside of school.

Consequently, for research purposes, child labor is any activity that threatens personal development and stops children from receiving an accurate education. Specifically, according to Peruvian law, every person until twelve-year-old is a child. Moreover, the Peruvian Instituto Nacional de Estadísticas e Informatica establishes that, on average, a worker's children labor between 14 and 19 hours a week (Instituto Nacional de Estadística e Informática, 2011).

4. Methodology

Logistic regression assumes that k is a set of independent observations $y_1, ..., y_k$, and the i-th term realizes the random variable Yi (Rodriguez, 2007). Moreover, Yi is a variable with a binomial distribution therefore:

$$Y_i \sim B(n_i, \pi_i) \tag{1}$$

Likewise, Rodriguez (2007) states that n_i is the probabilistic binomial denominator π_i. In addition, n_i has individual data for all i, which provides the stochastic model structure. Additionally, the logit probability is the predictors' linear function. Then:

$$logit(\pi_i) = x_i'\beta_j \tag{2}$$

Here x_i' is the covariance vector, while β is the logistic regression coefficient vector (Rodriguez, 2007). That equation becomes the definition of model structure. Since this model is linear with a binomial response and logit link (Rodríguez, 2007), it is more appropriate to take into account the distribution Y_i then the error distribution $Y_i - \mu_i$. Coefficient β and other regressions might have the same interpretation; however, they are logit values instead of averages. Consequently, β_j, is the logit probability change linked with a unit change in the *j-th* predictor while keeping all other predictors constant (Rodriguez, 2007).

Once the exponential affects equation two, it is feasible to find that *i-th* is:

$$\frac{\pi_i}{1-\pi_i} = \exp\{x_i'\beta\} \tag{3}$$

Rodriguez (2007) states that Eq. (3) defines a multiplicative model for probabilities. Hence, if the j-th predictor is changed by one unit and keeping the other variables constant, it is possible to multiply the chances with $\exp\{\beta_j\}$, then: $\exp\{x_i'\beta\}$ $\exp\{\beta_j\}$. Therefore, $\exp\{\beta_j\}$ represents the probability which is helpful to understand the variables' effect. Consequently, π_i probability is:

$$\pi_i = \frac{\exp\{x_i'\beta\}}{1+\exp\{x_i'\beta\}} \tag{4}$$

It is necessary to employ derivatives concerning x_j because the last equation does not give any predictors probability (Rodriguez, 2007). Then:

$$\frac{d\pi_i}{dx_{ij}} = \beta_j\pi_i(1-\pi_i) \tag{5}$$

Eq. (5) shows that the *j-th* predictor in probability π_i relies on coefficient β_j and probability value, which gives a better understanding of predictors' probability. Complementary analyses such as goodness of fit, sensitivity, and specificity are necessary to check the statistical accuracy of the Lasso proposal.

4.1. Lasso

The Least Absolute Shrinkage and Selection operator is a machine-learning approach that seeks the optimal quantity of independent variables (Fonti, 2017). For Benvenuto et al. (2018), Lasso is a supervised machine learning approach.

This method integrates wrapping and filtering methods (Fonti, 2017). Therefore, it can classify and erase coefficients (Tibshirani, 1996). Thus, Lasso can find a balance between variance and bias, allowing an optimal regression model without redundancy (Fonti, 2017). Equations six and seven provide the general form of Lasso:

$$Min = [(||Y - X\beta||_2^2)/n] \text{ subject to } \sum_{j=1}^k ||\beta||_1 < t, \tag{6}$$

or

$$\hat{\beta}(\lambda) = argmin_\beta(\frac{||Y-X\beta||_2^2}{n} + \lambda||\beta||_1). \tag{7}$$

In Eq. (7), $t >= 0$ reflects the upper limit of the sum of the coefficients. In equation 7, $|Y - X\beta|_2^2$ is like to $\sum_{i=0}^n (Y_i - (X\beta)_i)^2$; and $||\beta||_1$ is $\sum_{j=1}^k |\beta_j|$ y $\lambda >=0$. The λ –lambda- estimator manages the penalty force. Therefore, the penalty will have a direct correlation with λ. Moreover, λ and t have a negative relationship. Hence, when t goes to infinity, λ becomes 0. Only when the coefficient is zero does Lasso erases the variable.

Model selection of λ will be two Cross-Validation and Adaptative. Cross-Validation, or CV, begins with splitting the sample into training and testing sub-samples for obtaining a robust estimation (Reitermanová, 2010). Then, the CV splits the samples into ten folds. Once CV picks a fold, a linear regression employs the non-chosen ones (Stata, 2019). Thus, the regression coefficients estimate the selected fold.

Next, the technique gets the coefficients for the other folds. Therefore, the CV approach obtains ten average squared errors or MSE (Stata, 2019). The function stops when Lasso finds the λ minimum value. As a final step, the CV approach chooses the λ with the highest prediction power and the littlest MSE.

Meanwhile, the Adaptive Lasso or AV employs the previously described technique to select λ but in higher frequencies. Therefore, the AV model executes multiple functions simultaneously. Similar to CV, the AV function erases coefficients with zero values. However, the weak coefficients obtain penalty weights to make them become zero in the next step. Consequently, AV only chooses strong coefficients AV with the same methodology as the CV (Stata, 2019).

The way to select from CV or AV in logit regression is by looking at the deviance and deviance ratio. Therefore, the current study will choose the model with the lowest deviance or the highest deviance ratio.

In summary, the current study will employ the Lasso approach to find the predictors for the logit function. The study will analyze the selected variables using the previously described process. Results evaluation must ensure that the regression is accurate and provides correct information. Moreover, a marginal analysis will also be required to notice how each regressor contributes to the decision put in the dependent variables. The hanging and decision variables are as follows: a child studies and does not work, or a child works and does not study.

5. Results

Table 1
Initial variables

Independent variables	Dependent variable
Chief home gender	
Home chief education level	
Student gender	Children work
Academic Failure	
Geographic origin	
Native language	

Literacy

Poverty

Social stratification

Table 2
Lasso Results

Lasso model	ID	Lambda	Number of nonzero coefficients	Out/of/sample deviance ratio	Mean deviance
Cross validation	1	0.042	0	0.002	5.35E-01
	9	0.002	9	0.120	4.70E-01
	*10	0.002	9	0.120	4.70E-01
	11	0.001	9	0.120	4.70E-01
	14	0.000	10	0.119	4.71E-01
Adaptive	24	0.430	0	0.000	5.35E-01
	60	0.001	8	0.124	4.68E-01
	*61	0.001	8	0.124	4.68E-01
	62	0.001	8	0.124	4.68E-01
	99	0.000	8	0.124	4.68E-01

Table 3
Model selection

Lasso model	Sample	Deviance	Deviance ratio	Obs
Cross validation*	Training	0.461	0.137	2376
	Validation	0.503	0.132	1631
Adaptative	Training	0.461	0.137	2376
	Validation	0.504	0.132	1631

Table 4
Selected variables

Variables	Cross-validation	Adaptive
Home chief education level	√	√
Social stratification	√	√
Native language	√	√
House chief gender	√	√
Geographic origin	√	√
Academic failure	√	√
Gender	√	
Constant	√	√

Table 1 portrays the initial variables for the current study. The Lasso technique provides the best set of variables for building the algorithm. Then, according to the Cross-Validation technique, the id selected was 10, while 61 was for Adaptive. In Table 3, when comparing both methods, the Adaptive model was higher than the Cross-validation in deviance. Hence, the Cross-Validation proposal was the best for the study. Moreover, Table 3 shows the variables employed by each model.

Table 5
Logit regression

Variables	Probability	Standard error	z	p>z	95% confidence interval		chi2	p>chi2
House chief education level***	0.63	0.04	-7.97	0.00	0.56	0.71		
Social stratification***	2.24	0.21	8.49	0.00	1.86	2.70		
Native language***	1.22	0.04	5.55	0.00	1.13	1.30	0.1251	0.00
House chief gender**	1.46	0.20	2.80	0.01	1.12	1.89		
Geographic origin***	1.36	0.08	5.27	0.00	1.21	1.52		
Academic failure***	0.68	0.03	-7.85	0.00	0.61	0.75		

Gender**	1.26	0.12	2.45	0.01	1.05	1.51
Constant	0.00	0.00	-12.29	0.00	0.00	0.00

*** at 1%
** at 5%

Table 6
Post analysis

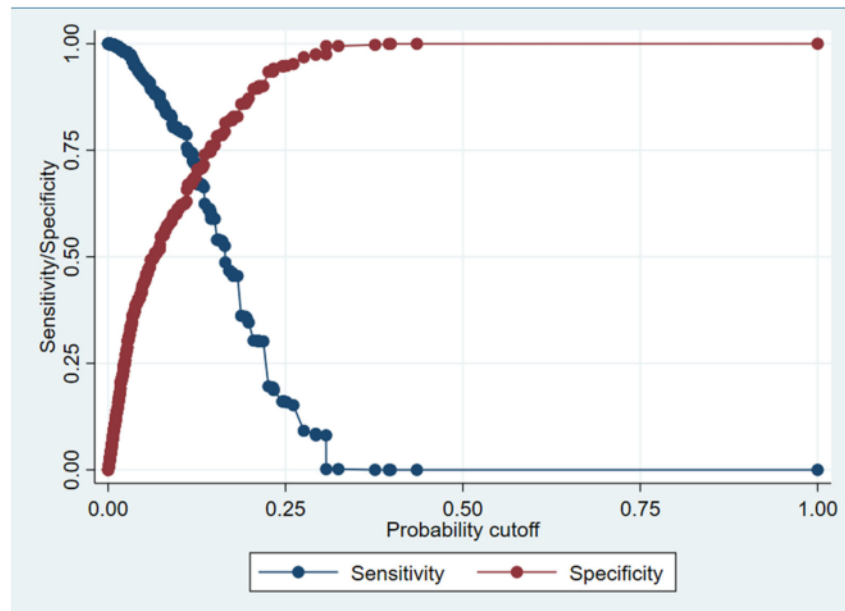| The goodness of fit test		
	Coefficient	p>chi2
GOF	14.68	0.0656
Classification		
Cut off	0.1191	
Sensitivity	72.66%	
Specificity	67.93%	
Correctly	68.39%	

Fig. 1. Sensitivity-Specificity chart

Table 5 shows that all variables selected by Lasso were statistically significant and that the model was well fitted. Moreover, Table 6 shows that the proposed model had rejected the null hypothesis, then there was evidence suggesting that the model was correct. Furthermore, both sensitivity and specificity were higher than 65%. Also, Fig. 1 provides a graphical view of the cut-off value, which was the point of balance between sensitivity and specificity.

It is important to note that all proposed variables were statistically significant. The logit regression shows that the higher education level of the home chief decreased the probability of child labor by about 0.63%. The student's social status also affected the likelihood of child labor. Here, it is necessary to clarify that the ENAHO dataset gave numbers inversely to the social class, then lower classes had higher numbers than higher ones. Therefore, lower-class students had 2.24% more involvement in child labor than upper-class students. The same happens with the native language. Students whose families taught them indigenous languages received higher numbers for classification purposes than students whose families spoke Spanish. Therefore, when the child comes from a family whose principal language is Quechua or another indigenous language, it has an extra probability of 1.22% to start working.

ENAHO pollsters gave males house chiefs a score of 1, while female ones received 0. Then, a child with a male as a house chief had a 1.46% more probability of working than students whose home chief was a female. Surprisingly, students living in the cities had 1.36% more chances of dropping out of school and working than students from rural areas. Furthermore, school failure provided a 0.68% probability for a child to work. Finally, the last variable showed that a man had 1.26% more chances to operate than a woman.

6. Discussion

The current study aimed to build an algorithm of child labor based on the ENAHO poll of students' social and economic characteristics. Consequently, the algorithm provided by Lasso gave a model where the variables portrayed in Table 5 had significant effects on child labor. The importance of the standard academic level described principally in the home chief was also observed in the research of Morales & Vargas (2019), Hernández & Vargas (2016), Khan et al. (2011), and Pariguana (2011). They all found that families with high academic levels do not allow their children to work quickly. On the other hand, some illiterate families do not recognize the importance of education for the future. Instead, they believe in working hard-working to achieve wellness.

Furthermore, the social roots of the children also influence the decision to stay at school or start working. Morales & Vargas (2019), Salata (2019), Khan et al. (2011), and Pariguana (2011) encountered that children coming from families with enough incomes to live did not leave school, at least for working reasons. Then, when the child's family has uncovered economic needs, the students start working, as found by Salata (2019) and Machado et al. (2015). The cultural aspects of the student family also played an essential role in child labor. Then students whose families had indigenous languages as their central idiom had more probability of starting working. In multicultural countries like Mexico and Bolivia, Morales & Vargas (2019) and Hernández & Vargas (2016) also discovered that children with indigenous roots had more academic difficulties than Spanish-speaking students. A possible explanation for that phenomenon might be that education policies segregated indigenous people for many years. Hence, education discrimination relegated indigenous people to poverty and illiteracy. Only in recent years have some Latin American governments had educational policies to integrate indigenous people into society. However, it is still not enough, and much work is to be done.

Furthermore, the gender of both parents and child plays a vital role in students' future. The current study found that students from houses with male chiefs had more probability of starting working than students from houses ruled by women (Pariguana, 2011). Machado et al. (2015) stated that men had more tendency to abandon school and start working than women. The current study found a similar event in the sample. A possible explanation can be the way the Peruvian society runs. Like many Latin American countries, the Peruvian culture has many characteristics of male domination. Then, male students have more pressure to start working to contribute to the family economy than women. This unfair situation provides women a slight advantage to stay at school and acquire primary education, which can be interrupted by pregnancy.

Interestingly, the study found that students from urban areas had more probability of starting working than children living in rural areas. This result does not match Alcázar's (2009) finding and Pariguana's (2011), who found that students from rural areas decided to work instead of studying because of idiomatic education barriers. A possible explanation here might be that rural children in the studied area do not have to abandon school to work since they help with familiar agricultural chores. In contrast, urban children face work journals that keep them from studying.

Also, school failure was a factor studied by Hernández & Vargas (2016) and Machado et al. (2015). In those studies and this one, school failure motivates children to start working rather than worrying about their academic performance. Also, their families can give them the idea that they do not serve to study, so they should focus all their efforts on work.

7. Conclusion

Child labor is a problem with many roots. Authors have studied this issue for years since it is a cause of low productivity for a country. Hence, it is not surprising that developing countries have high school dropout rates and child labor. Somehow, labor provides children with a sense of responsibility and satisfaction. However, labor becomes a problem when it prevents them from enjoying their age and studying. Social and economic reasons are behind the decision of a child to begin working.

According to the current analysis, the principal reasons are social class, gender, and where they live. The social environment can push or prevent children from labor. Educated families have more work opportunities and enjoy a better life standard than others. Then, they move their children to study rather than work at their educational age. Meanwhile, families with urgent economic needs and low academic levels push their child members to work. Furthermore, the role of gender portrays a scenario where women seem to appreciate the education value more than men. Again, the environment plays a vital role in child education since those living in urban areas have more probability of starting working than those in rural areas.

Therefore, how do we avoid child labor? First, people must value education. Society must see education as a powerful way to escape poverty and reach better living standards. Families whose child members work are not lazy or ignorant, but they do not believe that education in schools can help them with their urgent and future economic needs. Therefore, primary education must provide children with practical tools that help to change their family minds. Proposals like technical education should start as soon as possible in areas where child labor is high. Also, authorities can give social help to families with urgent economic needs. For instance, before the covid pandemic, the Peruvian government provided poor children with food. That measure avoided children being absent from school since they could eat for free at school. Additional measures like providing impoverished families with money for school attendance can prevent them from pushing their children to work. Other measures like psychological help and vocational assistance might be helpful to end someday child labor.

528

References

Abud, S. V. (2015). *Trabajo infantil : una realidad avasallante*. 8, 87–99.

Adam, S., Adom, D., & Bediako, A. B. (2016). The Major Factors That Influence Basic School Dropout in Rural Ghana: The Case of Asunafo South District in the Brong Ahafo Region of Ghana. *Journal of Education and Practice*, 7(28), 1–8. https://search.proquest.com/docview/1871574449?accountid=13042

Alcázar, L. (2009). *A d e s r p. 7.*

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4(3), S1–S39. http://www.jstor.org/stable/2534952

Benvenuto, F., Piana, M., Campi, C., & Massone, A. M. (2018). A Hybrid Supervised/Unsupervised Machine Learning Approach to Solar Flare Prediction. *The Astrophysical Journal*, 853(1), 90. https://doi.org/10.3847/1538-4357/aaa23c

Bureau of International Labor Affairs. (2019). *2018 findings of the worst forms of child labor: Required by the Trade and Development Act of 2000.* https://www.dol.gov/sites/dolgov/files/ILAB/child_labor_reports/tda2018/ChildLaborReportBook.pdf

Edmonds, E. V. (2015). *Economic Growth and Child Labor in Low Income Economies* (Issue 11).

Fonti, V. (2017). Feature Selection using LASSO. *VU Amsterdam*, 1–26. https://doi.org/10.1109/ACCESS.2017.2696365

Hernández, A. K., & Vargas, E. D. (2016). Condiciones del trabajo estudiantil urbano y abandono escolar en el nivel medio superior en México / Labor conditions of urban students and school dropout at the upper secondary level in Mexico. *Estudios Demográficos y Urbanos*, 31(3), 663. https://doi.org/10.24201/edu.v31i3.1653

Huebler, F. (2008). Child labor and school attendance: Evidence from MICS and DHS surveys. In *Understanding Children's Work Project Universidad*.

Instituto Nacional de Estadística e Informática. (2011). *Estado de la Niñez en el Perú*. Instituto Nacional de Estadística e Informática.

Khan, A., Azhar, M., & Shah, A. (2011). *Causes of Primary School Dropout Among Rural Girls in Pakistan, Sustainable Development Policy Institute (SDPI. July*, 119.

Machado, T., Ii, S., & Fernandes, S. (2015). *5 - Factors associated with dropout rates in public.pdf*. 757–772.

Miranda, F. (2018). Abandono escolar en educación media superior: conocimiento y aportaciones de política pública. *Sinéctica, Revista Electrónica de Educación*, 51. https://doi.org/10.31391/s2007-7033(2018)0051-010

Morales, J., & Vargas, Y. (2019). Determinantes De La Deserción Escolar Y El Trabajo Adolescente En Bolivia. *Investigacion & Desarrollo*, 18(2), 93–110. https://doi.org/10.23881/idupbo.018.2-6e

Pariguana, M. (2011). Trabajo adolescente y deserción escolar en el Perú. *Consorcio De Investigacion Económica Y Social*, 1(1), 1–49. http://old.cies.org.pe/files/documents/investigaciones/educacion/Trabajo-adolescente-y-desercion-escolar-en-el-peru.pdf

Reitermanová, Z. (2010). Data splitting. *WDS'10 Proceedings of Contributed Papers, Part I*, 31–36.

Revenberg, M. (2015). *Dropping out in a culture of poverty ? A case study on secondary and tertiary school dropouts in Curaçao. August*.

Rodriguez, G. (2007). Logit Models for Binary Data. In *Lecture Notes on Generalized Linear Models*. Princeton University. https://doi.org/10.1002/ird.2208

Ruiz, I. (2019). *América del Sur : avances y retos hacia el cumplimiento de la Meta 8 . 7 de los ODS Informe del Mapeo sobre la problemática del trabajo infantil y*.

Salata, A. (2019). Razões da evasão: abandono escolar entre jovens no Brasil. *Interseções: Revista de Estudos Interdisciplinares*, 21(1), 99–128. https://doi.org/10.12957/irei.2019.42305

Skoufias, E. (2005). *PROGRESA and Its Impacts on the Welfare*.

Stata. (2019). *Stata Lasso Reference Manual Release 16*. https://www.stata.com/manuals/lasso.pdf

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society*, 58(1), 267–288.

Uddin, M. N., Hamiduzzaman, M., & Gunter, B. G. (2011). Physical and Psychological Implications of Risky Child Labor: A Study in Sylhet City, Bangladesh. *SSRN Electronic Journal*, July. https://doi.org/10.2139/ssrn.1428206

Velazco, J. (OIT). (2016). *Magnitud y Características del Trabajo Infantil en Perú. Informe 2015 - Análisis de la Encuesta Nacional de Hogares (ENAHO) y de la Encuesta sobre Trabajo Infantil (ETI)*. http://white.lim.ilo.org/ipec/documentos/informeti_2015_peru.pdf