



Utilizing CART Machine-Learning Models to Explore the Intersections of Determinants of Key Early Childhood Development (ECD) Outcomes

Contributors:

Bilal Afroz, Senior Consultant – MLE

Dr. Francis Xavier Rathinam, Associate Director & Practice Lead – MLE

Introduction

Early childhood is a critical period that affects the holistic development of an individual and determines their ability to reach their optimum health, social, and economic potential. This period shapes one’s adult life, as the cognitive and social skills of a child begin to develop and mature during this period.¹

Evidence from literature² suggests that children from low-income families are more likely to exhibit developmental delays, behavioural problems, and other disabilities than children from wealthier families. Children from vulnerable and disadvantaged families are most likely to miss the developmental milestones in their early childhood due to their extreme exposure to the cumulative effects of risk factors. These risk factors can include a lack of: basic water, sanitation, and hygiene

(WASH) services; quality health services; nutritional inputs; and quality daycare centres and preschools. As a result of these risks, these children are less likely to be enrolled in schools at the right age, which further affects their development.³

Nutrition is another important indicator of children’s overall health. Children are considered “well-nourished” if they have access to sufficient and diverse foods and are not exposed to repeated illnesses, such as parasitic infections, that prohibit the absorption of essential nutrients. On the contrary, malnutrition contributes to poor health and poor development in young children. Malnutrition is a major contributor to child mortality globally and is estimated to be associated with a third of deaths under the age of five.⁴ Undernourished children who survive are more likely to suffer from common illnesses and are less likely to reach their growth and developmental potential.^{5, 6} Furthermore, boys are

¹ <https://www.unicef.org/early-childhood-development>

² https://apps.who.int/iris/bitstream/handle/10665/97942/9789241503549_eng.pdf?sequence=1&ua=1

³ <https://openknowledge.worldbank.org/bitstream/handle/10986/9916/570760BRI0Find10Box353745B01PUBLIC1.pdf?sequence=1&isAllowed=y>

⁴ Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, Mathers C, Rivera J. Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet*. 2008;5-22.

⁵ Schaible UE, Kaufmann SHE. Malnutrition and Infection: Complex Mechanisms and Global Impact. *PLoS Medicine*. 2007;4(5):e115.

⁶ Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B, the International Child Development Steering G.

usually more stunted than girls, which is a pattern seen across Sub-Saharan Africa.⁷

Evidence also suggests that a positive family environment is critical for optimal child development. Exposure to household conflict, violent practices (psychological aggression or physical punishment), and inadequate care of children at home have been shown to have negative consequences on children's development and later life outcomes.^{8,9}

These determinants of early child development, such as wealth, the child's gender, education levels of caregivers, violent practices¹⁰ against children, inadequate care¹¹ of children at home, and the mental health of the primary caregiver, have the most significant impact on a child's development.¹²

In this context, understanding the intersections of these determinants of key early child development outcomes, such as stunting, Early Childhood Development (ECD) Index, and Observation of Mother-Child Interactions (OMCI) scores, has great potential in maximizing the efficiency of the development programs promoting early childhood development by facilitating a more precise targeting of the sub-groups in greatest need. However, the understanding of intersections of determinants of key early child development outcomes is still evolving and is limited by inadequacies of methods to quantitatively understand the intersections at play.

In this brief, we share our experience of applying machine-learning models in exploring determinants of key early childhood development outcomes from an end-line evaluation Athena Infonomics conducted for UNICEF Rwanda. The final report can be accessed [here](#). The end-line evaluation was conducted with 882 households in ten districts of Rwanda, where UNICEF implemented the "Early Childhood Development and Family (ECD&F)" program between 2014 and 2020. The evaluation team used *Classification and Regression Tree* (CART) prediction models to explore intersections of determinants of three key child development outcomes: stunting; the

Child Development Index; and Observation of Mother-Child Interactions (OMCI) scores.

Classification and Regression Trees (CART)

The *Classification and Regression Tree* (CART) algorithm, introduced by Breiman et al.,¹³ was used for this analysis. CART is a non-parametric, statistical approach that uses the recursive partitioning technique to split a sample population into sub-groups based on a predefined criterion. When provided with a dependent variable (either categorical or continuous) and a set of independent demographic variables hypothesized to influence the dependent variable, CART creates mutually exclusive sub-groups based on combinations of demographics within the sample and the proportion of individuals in a particular sub-group who are likely to engage in the behaviour/practice represented by the dependent variable.¹⁴

Dependent Variables

Three key measures related to child development were considered: stunting; children between 36 and 59 months who were considered "developmentally on track" in the ECD Index; and the total Observation of Mother-Child Interactions (OMCI) score.

Independent Variables

Demographic indicators, such as household wealth, the primary caregiver's highest education, the age group of children, and the gender of the child, were included as independent variables. In addition to these demographic indicators, factors generally known to affect child development outcomes—such as violent practices against children, inadequate care for children in homes, support for learning at homes, decision making related to what a child eats, children's books at home, and the mental health of the primary caregiver—were included as independent variables. It should be noted that the

Developmental potential in the first 5 years for children in developing countries. *Lancet*. 2007;369(9555):60-70.

⁷ Wamani H, Åström AN, Peterson S, Tumwine JK, Tylleskär T. Boys are more stunted than girls in Sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys. *BMC Pediatr*. 2007;7(17).

⁸ Schilling, Elizabeth A., Robert H. Alseltnine Jr. and Susan Gore, 'Adverse Childhood Experiences and Mental Health in Young Adults: A longitudinal survey, *BMC Public Health*, vol. 7, no. 30, 2007, pp. 1–10.

⁹ Walker, Susan P., et al., 'Child Development: Risk factors for adverse outcomes in developing countries', *The Lancet*, vol. 369, no. 9556, 2007, pp. 147–157.

¹⁰ exposure to either psychological aggression or physical punishment

¹¹ Inadequate care is defined as a child who is left alone or in the care of another child less than 10 years of age for an hour or more in the week

¹² <https://www.who.int/publications/i/item/early-child-development-a-powerful-equalizer-final-report-for-the-world-health-organization-s-commission-on-the-social-determinants-of-health>

¹³ Breiman L, Friedman J, Olshen R, Stone C. *Classification and regression trees*. Monterey, Calif., USA: Wadsworth, Inc; 1984.

¹⁴ Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of behavioral medicine*. 2003; 26(3):172±81. https://doi.org/10.1207/S15324796ABM2603_02 PMID: 14644693

majority of the primary caregivers in the end-line evaluation were mothers (85%), so we could not include the gender of the primary caregiver as an independent variable in the CART model.

Discussion of Results

a) Stunting in Children Between 24 and 59 Months

Stunting is the impaired growth and development that children experience from poor nutrition, repeated infection, and inadequate psychosocial stimulation. Children are defined as stunted if their height-for-age is more than two standard deviations below the WHO Child Growth Standards median.¹⁵

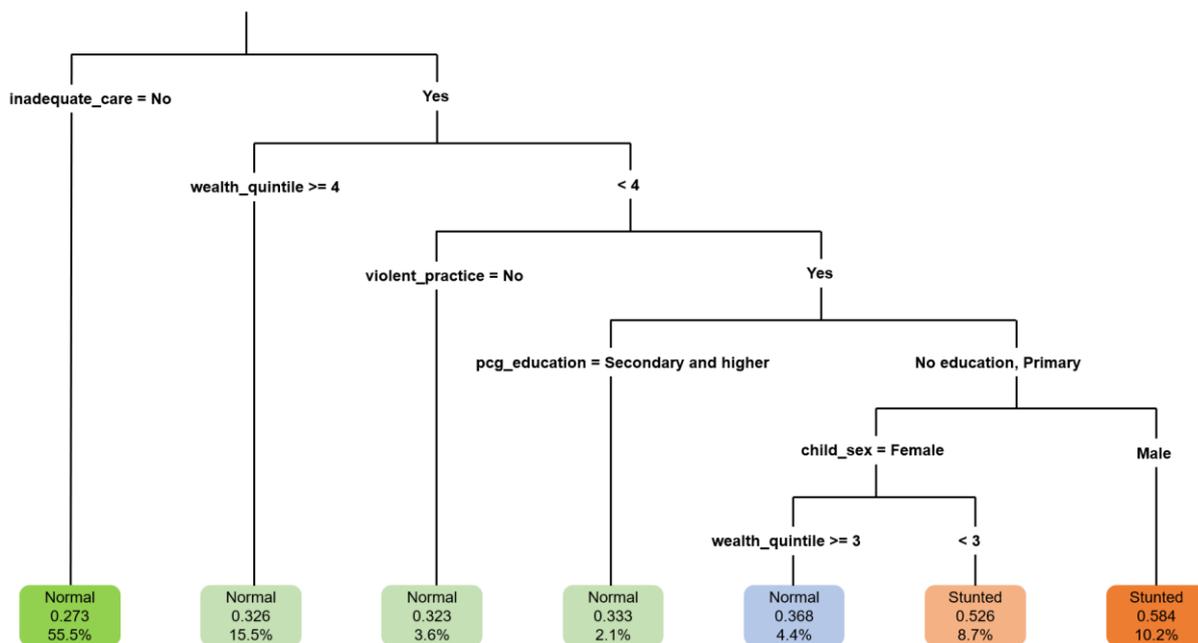


Figure 1. CART model predicting the stunting rates in children between 24 and 59 months, based on initial conditions.

The model in Figure 1 shows that the children between 24 and 59 months who were not exposed to any inadequate care had relatively lower stunting rates (27.3 percent), indicating the key role of household behaviour dynamics on child nutrition outcomes. The stunting rates were quite high (58.4 percent) among the group of children between 24 and 59 months who were exposed to any inadequate care, belonged to lower wealth quintiles (1, 2, 3), were exposed to any violent practices, whose caregiver had no education or primary education, and were male. This shows the compounding effect of various determinants on stunting rates. Even among the households where the children were exposed to any inadequate care, the households

belonging to higher wealth quintiles (4, 5) had relatively low stunting rates (32.6 percent).

Furthermore, as can be observed in the model, children who were not exposed to any violent discipline, whose primary caregiver had secondary and higher education, and households belonged to higher wealth quintiles were less likely to be stunted.

b) Children “Developmentally on Track” in the ECD Index

The measurement of children’s development in Rwanda, as in other low- and middle-income countries, is challenging due to the lack of locally validated measures of children’s development.^{16, 17}

¹⁵ <https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell>

¹⁶ Gladstone M.J., et al., ‘Can Western Developmental Screening Tools be Modified for Use in a Rural Malawian Setting?’ *Archives of Disease in Childhood*, vol. 93, no. 1, 2007, pp. 23–29.

¹⁷ Yousafzai, Aisha K., Paul Lynch and Melissa Gladstone, ‘Moving beyond Prevalence Studies: Screening and interventions for children with disabilities in low-income and middle-income countries’, *Archives of Disease in Childhood*, vol. 99, no. 9, 2014.

To assess children’s development, the ECD Index¹⁸ was included at the end-line, which was developed specifically for use in national-level surveys to evaluate the development of children between 36 and 59 months of age and was first used in Rwanda as part of the 2014–2015 DHS.¹⁹

The ECD Index examines four domains of development: literacy and numeracy, and physical,

social-emotional, and learning among children. Mothers or primary caregivers are asked 20 questions about the way their children behave in certain everyday situations, and the skills and knowledge they have acquired. Children between the ages of 36 and 59 months are considered “developmentally on track” if they are on track in at least three of these four domains.

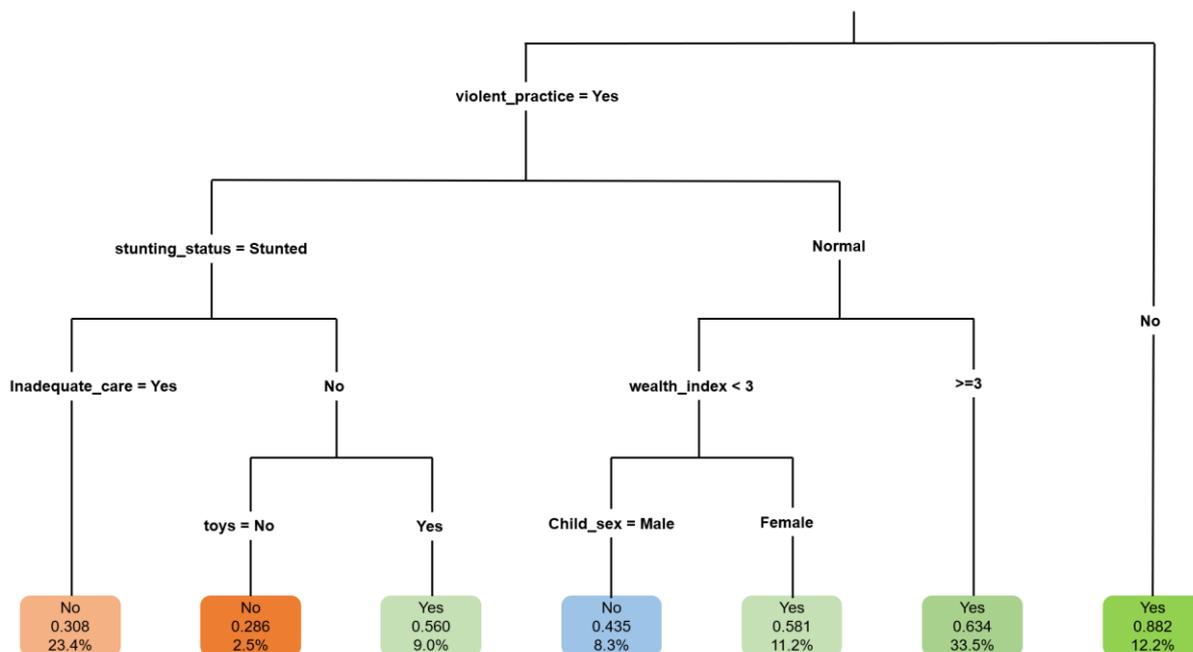


Figure 2. CART model predicting if a child is “developmentally on track” on the ECD Index, based on initial conditions.

The model in Figure 2 shows that the children between 36 and 59 months who were not exposed to any violent discipline were relatively more likely to be developmentally on track (88.2 percent). This shows the possible impact of household behaviour dynamics on the development status of the child. Furthermore, the children who were exposed to any violent practices but who were receiving adequate nutrition and who belonged to a higher wealth quintile (3, 4, 5) were more likely to be developmentally on track (63.4 percent). Likewise, female children who were exposed to any violent practices but who were receiving adequate nutrition, and belonged to the lower wealth quintile (1, 2) were

more likely to be developmentally on track (58.1 percent) when compared to male children (43.5 percent). In addition, children who were stunted, who

experienced any violent discipline but received adequate care and had toys in the home were more likely to be developmentally on track (56 percent) when compared to children who were stunted, who experienced any violent discipline, received any inadequate care, and did not have toys in the home (28.6 percent). This indicates the critical role of the nutrition status of the child, adequate care practices, and the availability of toys at home on the development status and well-being of the child.

c) Total OMCI Score (Mother and Child)

Observation of Mother-Child interactions (OMCI) is a measure where the primary caregiver is provided with a children's book in the local language and asked to play with their children for five minutes. This interaction is observed and scored by a trained observer. Rasheed and Yousafzai (2015) developed

¹⁸ Kariger, Patricia, et al., ‘Indicators of Family Care for Development for Use in Multicountry Surveys’, *Journal of Health, Population and Nutrition*, vol. 30, no. 4, 2012, pp. 472–486.

¹⁹ National Institute of Statistics, ‘Rwanda Demographic and Health Survey 2014–15 Final Report’, Republic of Rwanda, Kigali, 2016.

the Observation of Mother-Child Interaction (OMCI) tool to capture reliable and valid measurements of parental sensitivity and responsivity.

The tool yields an overall interaction score, along with separate scores for the parent and child. The

total OMCI score is a sum of 19 items,²⁰ with higher scores indicating a greater frequency of behaviours associated with better quality and more responsive interactions. The maximum possible score on the OMCI is 57.

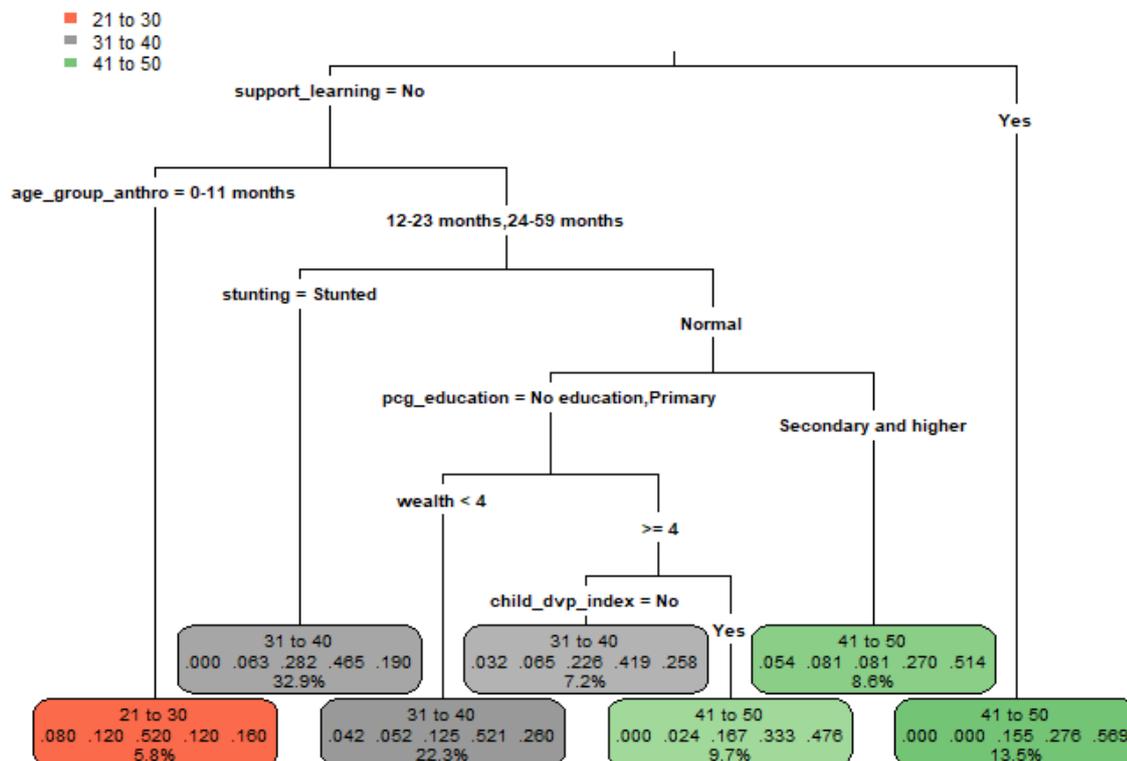


Figure 3. CART model predicting the total OMCI score, based on initial conditions.

The model in Figure 3 shows that the total OMCI score among households where caregivers engaged in three activities to support learning is likely to be higher (41 to 50), which signifies the importance of household behaviour dynamics on the development of the child. The model shows that the households with children from the younger age group (0-11 months) tend to have a relatively lower total OMCI score (21 to 30) when compared to the older group (12-23 months and 24-59 months). Furthermore, among the older children (12-23 months and 24-59 months) whose caregivers did not provide learning support and who were stunted, the total OMCI score

was likely to be on the lower side (31 to 40). Interestingly, even if the caregivers did not engage proactively in providing learning support to children,

older children (12-23 months and 24-59 months) who were not stunted (normal) and whose caregivers had secondary and higher education were likely to have a high total OMCI score (41 to 50). As can be observed in the figure, older children (12-23 months and 24-59 months) who belong to the lowest wealth quintiles (1, 2, 3) and who were not developmentally on track on the ECD Index were more likely to have low total OMCI scores.

Conclusion

The CART prediction models provide important insights into the intersections of key determinants in predicting key early childhood development outcomes (stunting, ECD Index, and OMCI scores). The analysis reveals the role of interactions between exposure to inadequate care and violent practices, wealth, and literacy levels of primary caregivers in

²⁰ The total OMCI score is a sum of 19 items covering maternal affect, maternal touch, maternal verbalization, sensitivity and contingent responses, scaffolding, language stimulation, focus,

child affect, child focus, child's communication efforts, and mutual enjoyment.

creating substantial disparities in early childhood development outcomes among children under the age of five living in Rwanda. We also found that the gender of the child and the availability of toys at home interact with inadequate care and violent practices, wealth, and literacy to exacerbate disparities in early childhood outcomes. The understanding of such interactions is still evolving and further research across multiple populations and geographies is required to determine its broader application. Future research should assess the utility

of exploring intersections as an approach to targeting the right combination of interventions to bring about better development outcomes for children below five years of age. In this context, the application of machine-learning techniques may prove to be a useful tool for exploring intersections of determinants of early childhood development outcomes and may be utilized by potential interventions that seek to optimize child development.