Technical Audit of an Automated Decision System: Costa Rican Household Poverty Prediction

Group #9: Yui Cao (yc4234) & Kelly Lin (xl3309)

1. Background

The chosen ADS was a solution from the Kaggle competition "Costa Rican Household Poverty Prediction." This competition, launched by IDB (Inter-American Development Bank - the largest source of development financing for Latin America and the Caribbean), was to develop an efficient model that accurately predicts poverty levels using PMT (Proxy Means Test, a model that uses household or individual characteristics) based on a dataset of Costa Rican household and individual characteristics. Developing an accurate ADS will help IDB and other public agencies to address households in need of financial assistance and accommodate appropriate social programs. This competition's objective is to predict poverty on a household level, with the given individual level dataset with unique features about the individual and information about their household. We suspect that if the model aggregates the individual features on the household level, the model may fail to address certain sub-populations, failing to address the needs of sub-populations under poverty. Furthermore, we will look in detail at the potential trade-offs between the increase in performance and fairness. We will compare the accuracy, fairness metrics, and robustness to evaluate the effectiveness of the ADS developed by the author.

2. Input and Output

2.1 Data Used

The 2 datasets in ADS include individual and household-level data from Costa Rica, provided by IDB, featuring a range of ordinal, binary, continuous, and categorical variables that detail demographics, socio-economic status, and living conditions, collected through detailed household surveys. The competition aimed to reflect diverse influencing poverty, ensuring the model developed could generalize across different household environments. The competition provides a training set with 9557 rows and 143 columns, and a testing set with 23856 rows and 142 columns where each row represents an individual and column represents an individual or household feature. The training set includes a 'Target' column, categorizing poverty levels from 1 (extreme poverty) to 4 (non-vulnerable). Test set does not include Target label, since this was a Kaggle competition, and test set was used for evaluation purposes. The author validates the model by using cross-validation with training set.

2.2 Input Features

2.2.1 Data Types (key features selected)

Variables	Description	Data Type
Area1	urban	Categorical
Area2	rural	Categorical
idhogar	household level identifier	String
v2a1	monthly rent payment	Continuous/Numerical
hacdor	overcrowding by bedrooms	Binary

rooms	number of all rooms in the house	Discrete/Numerical
hacapo	overcrowding by rooms	Binary
v14a	has bathroom in the household	Binary
refrig	if the household has refrigerator	Binary
v18q1	number of tablets household owns	Discrete/Numerical
r4h3	total males in the household	Discrete/Numerical
r4m3	total females in the household	Discrete/Numerical
meaneduc	avg. education level of adults in the household	Continuous/Numerical
escolari	years of schooling	Discrete/Numerical
rez_esc	years behind in school	Discrete/Numerical
hhsize	household size	Discrete/Numerical

Table 1: Selection of Important Features

2.2.2 Missing Values (Imputation discussed in 3.1)

- **rez_esc** (years behind in school) has 27,581 missing values (82.5%). For the families with a null value, is possible that they have no children currently in school.
- **v18q1** (Number of tablets household owns) has 25,468 missing values (76.2%). Missing values likely indicate that no tablets are owned by the household. The data collectors may have left these fields blank when no tablets were present.
- **v2a1** (Monthly rent payment) has 24,263 missing values (72.6%). Compared with the distribution of **tipovivi** (the ownership/renting status of the home), we know that the households that do not have a monthly rent payment generally own their own home.

2.2.3 Pairwise Correlations

To understand data interdependencies affecting the model's decisions, we selected key features by analyzing their correlations with Target variable and crucial demographic information, considering the large number of features available. He used Spearman correlation for ordinal variables and Pearson correlation for numerical features' correlation with poverty class.

```
Most negative Spearman correlations:
       feature     scorr       pvalue
97     warning -0.307326  4.682829e-66
68  dependency -0.281516  2.792620e-55
85   hogar_nin -0.236225  5.567218e-39
80        r4t1 -0.219226  1.112230e-33
49       eviv1 -0.217803  2.952571e-33

Most positive Spearman correlations:
             feature     scorr       pvalue
23         cielorazo  0.300996  2.611808e-63
95             floor  0.309638  4.466091e-67
99  phones-per-capita  0.337377  4.760104e-80
96   walls+roof+floor  0.338791  9.539346e-81
0              Target  1.000000  0.000000e+00
```

Figure 1: Spearman Correlation

```
Most negatively correlated variables:
        feature     pcorr
0       warning -0.301791
1     hogar_nin -0.266309
2          r4t1 -0.260917
3  overcrowding -0.234954
4         eviv1 -0.217908

Most positively correlated variables:
             feature     pcorr
95  phones-per-capita  0.299026
96              floor  0.307605
97   walls+roof+floor  0.332446
98           meaneduc  0.333652
99             Target  1.000000
```

Figure 2: Pearson Correlation

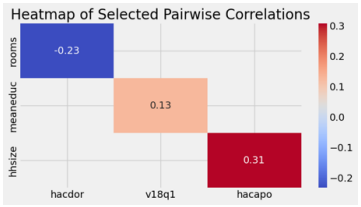

Figure 3: Heatmap of Selected Pairwise Correlation

Based on highly correlated features with Target, we identified key features that represent demographic characteristics of the household, such as dependency, education level of male vs. female head, gender ratios in the household, and others like rooms vs. overcrowding (hacdor), education level vs. tablet ownership, and household size vs. overcrowding by rooms. A pairwise heatmap was generated, highlighting:

- **Spatial and Living Conditions:**
  - **(rooms) Rooms vs. (hacdor) Overcrowding by Bedrooms:** More rooms correlate with less bedroom overcrowding, indicating spacious homes have less congestion.

- **(hhsize) Household Size vs. (hacapo) Room Overcrowding:** Larger household has more room overcrowding, indicating spatial challenges with increasing family size.
- **(hhsize) Household Size vs. Female Counts: r4m1** (Females younger than 12), **r4m3** (total females), and household size are highly correlated, indicating that increases in female counts are associated with larger household sizes.

- **Education Access and Technological Availability:** A weak positive correlation exists between **meaneduc** (adult education levels) and **v18q1** (tablet ownership), implying that higher education may improve access to technology.
- **Gender Dynamics and Educational Attainment:**
  - **(edjefe) Male vs. (edjefa) Female Education on (Target) Poverty:** Male education has a stronger correlation with poverty than female education, highlighting different impacts based on gender.
  - **(r4m3) Total Female Count vs. (meaneduc) Education:** More females in a household correlate with higher education but also with greater poverty vulnerability.
  - **(r4h3) Total Males Count vs. (meaneduc) Education:** It's negatively correlated, showing lower education levels in male-dominant households.
- **Overcrowding vs. Poverty and Warning:** Overcrowding is correlated with both the poverty level and warning, indicating that higher levels of overcrowding are associated with increased poverty and greater concerns about living conditions.

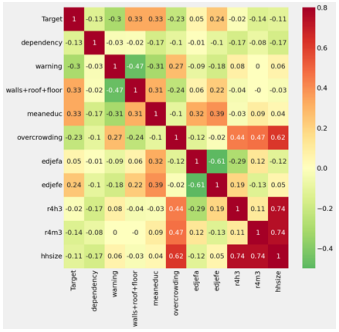

Figure 4: Heatmaps of Selected Pairwise Correlations

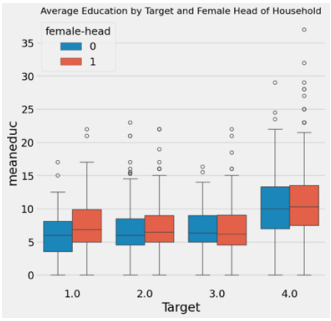

Figure 5: Avg. Education by Target and Female Head of Household

Upon analyzing meaneduc distribution across 4 poverty classes by female heads, it was found females heads have higher education levels but are more often classified as vulnerable.

2.3 Output

The system classifies households into one of four poverty levels, labeled as classes 1 to 4, which help guide interventions and allocate resources based on the poverty status of each household. Each class label indicates a distinct poverty status: **1 = extreme poverty, 2 = moderate poverty, 3 = vulnerable households, and 4 = non-vulnerable households.** However, a key issue here is the lack of data for vulnerable households, leading to an imbalance that causes models to overfit on non-vulnerable households, thus failing to accurately identify those needing financial assistance. Additionally, female household heads are underrepresented across all poverty levels, especially in the non-vulnerable class.

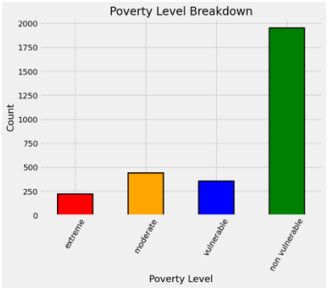

Figure 6: Poverty Level Breakdown

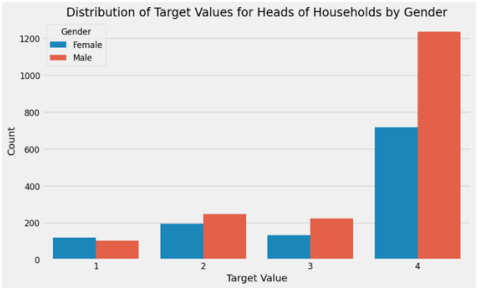

Figure 7: Distribution of Target Values for Heads of Households by Gender

3. Implementation and Validation

3.1 Data Cleaning & Pre-Processing

To make operations easier, the author joined the training and testing dataframes together so he can apply the same operations to both dataframes and get the same features in the end.

- **Mixed Data Types:** In object columns, the **dependency**, **edjefe** (years of education of male head of household), and **edjefa** (years of education of female head of household) have a mix of strings and numbers. Based on the documentation ("yes" = 1 & "No" = 0), the author corrected the variables using a mapping and converted them to floats.
- **Missing Variables:** The author imputes missing values based on the available information of that variable and other redundant or related variables.
    - **rez_esc** (the years behind in school): defined only for ages 7-19. The author imputes 0 for ages outside of this range and adds a Boolean flag column **rez_esc-missing** for other cases. Any value outlier is capped at 5.
    - **v18q1** (number of tablets household owns): cross-referenced with **v18q** (if a family owns a tablet), missing values in **v18q1** are imputed 0, indicating no tablet ownership.
    - **v2a1** (monthly rent payment): the absence of rent payment usually indicates homeownership. Missing values are imputed as 0 for homeowners and a flag column **v2a1-missing** is added to track households with missing values, noting significant numbers of households still show missing rent data.

3.2 Feature Engineering

The author used sklearn's RFECV for feature selection. However, due to version differences in sklearn, using the original code to replicate selected features resulted in discrepancies in the number of features chosen. We confirmed that all steps prior to the feature (data cleaning, pre-processing, aggregations) were consistent and that the training datasets were identical before applying feature selection. We couldn't resolve the problem entirely, so we extracted the selected features directly from the original solution (refer to our Colab for details). After verifying that both datasets were identical post-feature selection, we proceeded with model development.

With the focus on household-level poverty prediction, he aggregates the individual level features into household-level data: 1) Break variables into household and individual levels, 2) Find

suitable aggregations for the individual data (using statistical methods for ordinal variables and limited statistics for Boolean variables), 3) Join these aggregations to the household-level data.

- **Squared Variables:** The author removes squared variables due to their high correlation with non-squared versions, which are sufficient for model development, as squared variables can negatively impact more complex models.
- **Household Level Variables:** He subsets to the head of household and household level variable for prediction, keeping most of the variables as it while removing redundant ones and adding derived features from existing data.
    - **Redundant Variables**: He removes redundant variables that are highly correlated, using a threshold of 0.95. Variables like **tamhog** (household size), **hogar_total** (total individuals in the household), **r4t3** (total people in the household), which all indicate household size, are removed due to redundancy and high correlation. Also, **area2**, indicating a rural zone, is removed as it is redundant with **area1**, indicating an urban zone. The author creates a new variable **hhsize-diff**, to capture discrepancies between the number of people and the household size, addressing observed mismatches.
    - **Creating Ordinal Variable:** For household construction conditions, the author turns them into ordinal variables reflecting the inherent order (bad < regular < good) using np.argmax and drops the original redundant variables for roof (etecho1: bad, etecho2: regular, etecho3: good). A new feature, **walls+roof+floor**, is created, summing these features to provide an overall measure of the house's structural quality.
    - **Warning + Bonus Feature:** He created two new features for predicting household poverty: **warning** (deducts a point for each missing basic amenity like a toilet or electricity) and **bonus** (awards points for having items like a refrigerator or tablet).
    - **Per Capita Feature:** The author calculated per-capita metrics including phones-per-capita, tablets-per-capita, rooms-per-capita, and rent-per-capita.
- **Individual Level Variables:** The author aggregates individual data by grouping by **idhogar** (family ID). Redundant columns are dropped, including one of each pair of variables with a correlation over 0.95. **Male** is removed due to existing **female** features.

3.3 LightGBM with Hyperparameter Tuning

The author implemented Light GBM using 5-fold stratified cross-validation and early stopping to prevent overfitting, enhancing model training speed and reducing bias compared to single-fold methods. Overall, early stopping with cross-validation is the best method to select the number of estimators in the Gradient Boosting Machine. The author set the n_estimators to 10000 and the hyperparameters were determined from the author's previous successful works. **The 5 cross-validation score is 0.41963** with a standard deviation of 0.0133. The most influential features identified were related to age and education.

Further testing with features selected via recursive feature elimination resulted in an **improved cross-validation score of 0.42704** with a standard deviation of 0.01953. The final model, using

10-folds cross-validation on these features, **achieved the best score of 0.42736**. This model, while not fully optimized at submission, showed potential for further performance gains.

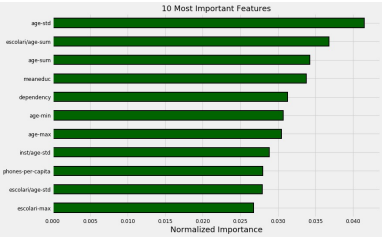

Figure 8: 10 Most Important Features

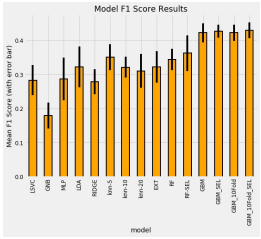

Figure 9: Model F1 Score Results

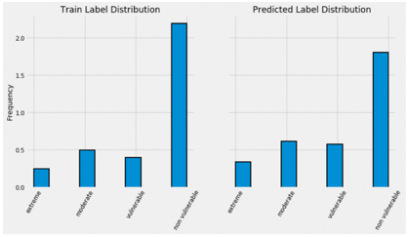

Figure 10: Train & Predicted Label Distribution

4. LightGBM Outcome

Here, we faced challenges replicating the original model from Kaggle, primarily due to version discrepancies. The latest LightGBM version does not support "dart" boosting mode with early stopping, so we switched to "gdb" to enable early stopping and optimize hyperparameters. We retained the original model's predefined hyperparameters, with the only change being the boosting type. Despite slight variations in feature importance rankings, our model largely mirrors the original in terms of F1 score. For further details, please refer to our Colab.

The author visualized the label distributions of the training set and predictions (Figure 10 above), revealing that the predicted outcomes closely mirror the original label distribution, with the largest group being nonvulnerable, followed by moderate, vulnerable, and extreme poverty. It indicates the model's success in replicating the target label distribution from original population.

4.1 Accuracy

4.1.1 F1 Scores

- **Overall:** The competition uses the Macro F1 score for evaluation, averaging the F1 scores across four classes (extreme, moderate, vulnerable, nonvulnerable) without considering label imbalances. The Light GBM model achieved a **macro F1 score of 0.42712**, closely matching the original solution's 0.42704. Despite being low, this score is deemed reasonable given the significant imbalance favoring non-vulnerable households. Further insights were gained by examining the F1 scores for each class.
- **By poverty level:** The F1 score breakdown shows the model performs best at predicting non-vulnerable households, likely due to class imbalance, with a **score of 0.79 for class 4**, while scores for the other classes are around 0.2-0.3. This skew towards nonvulnerable predictions is problematic for a system intended to identify households needing financial aid, as it fails to accurately predict vulnerable groups.
- **By poverty level and gender:**

Table 3: Classification Report for Female

Class	Precision	Recall	F1-Score	Support
1	0.34	0.46	0.39	46
2	0.32	0.36	0.34	67
3	0.18	0.23	0.20	48
4	0.81	0.70	0.75	256

Table 2: Classification Report for Female

Table 4: Classification Report for Male

Class	Precision	Recall	F1-Score	Support
1	0.09	0.06	0.07	34
2	0.37	0.44	0.40	72
3	0.25	0.38	0.30	65
4	0.85	0.77	0.81	412

Table 3: Classification Report for Male

During the EDA process, we observed variations in education levels and poverty across genders, leading us to further dissect the F1 scores by gender and poverty level.

- **For Females - Class 1 (F1-Score: 0.39):** This class has a moderate F1 score, which indicates a reasonable balance between precision and recall. However, the model struggles somewhat with correctly identifying this class as the recall (0.46) is less than ideal; **Class 2 (F1-Score: 0.34):** Similar to Class 1, the F1 score here is also moderate but slightly lower, suggesting issues with both precision and recall; **Class 3 (F1-Score: 0.20):** This score is low, highlighting significant difficulties in predicting this class accurately. Both precision (0.18) and recall (0.23) are poor, suggesting that improvements are needed for this class; **Class 4 (F1-Score: 0.75):** The highest F1 score among the female classes indicates good model performance for this class, with relatively high precision and recall.
- **For Males - Class 1 (F1-Score: 0.07):** Very low F1 score, indicating very poor model performance for this class with extremely low precision (0.09) and recall (0.06); **Class 2 (F1-Score: 0.40):** Moderate performance, with better recall than precision. This suggests that while the model is identifying relevant cases more frequently, it is still making a considerable number of mistakes; **Class 3 (F1-Score: 0.30):** Slightly below moderate, this score suggests some capability to identify this class but with room for improvement in both precision and recall; **Class 4 (F1-Score: 0.81):** Excellent performance, with high precision and recall, indicating that the model performs well in identifying this class.

4.1.2 ROC & AUC

We used a one-vs-rest method to generate ROC curves to gauge model accuracy. The micro-average ROC curve, which compiles all classification outcomes across classes, shows an **AUC of 0.84**, indicating good overall performance in predicting TP for each class. The macro-average ROC curve, which calculates ROC metrics for each class independently before averaging, ensures equal treatment of all classes. Analysis of ROC curves by class revealed that while class 4 performs best with an **AUC of 0.83**, other classes underperform, with class 3 near a random classification level at an AUC of 0.69. This underscores the model's difficulty in differentiating among poverty levels. Additionally, a low macro F1 score and moderate ROC AUG suggest that although the model ranks poverty effectively, it struggles to precisely classify households within each level, particularly failing to distinguish between poverty classes 1-3. The model is more accurate in predicting non-vulnerable households but less so for vulnerable ones.

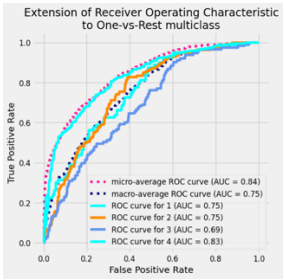

Figure 11: ROC Curves

Table 1: Overall Metrics by Gender

Metric	Class	Female	Male
Precision	1	0.3443	0.0909
	2	0.3243	0.3678
	3	0.1803	0.2500
	4	0.8145	0.8529
Recall	1	0.4565	0.0588
	2	0.3582	0.4444
	3	0.2292	0.3846
	4	0.7031	0.7743
FPR	1	0.1078	0.0364
	2	0.1429	0.1076
	3	0.1355	0.1448
	4	0.2547	0.3216
FNR	1	0.5435	0.9412
	2	0.6418	0.5556
	3	0.7708	0.6154
	4	0.2969	0.2257
Demographic Parity Ratio		0.8261	
Class 4 Selection Rate		0.5300	0.6415

Table 4: Overall Metrics by Gender

Table 2: Overall Metrics for Class 4 by Gender

Metric	Male	Female
Precision	0.8145	0.8529
Recall	0.7031	0.7743
False Positive Rate	0.9318	2.3913
False Negative Rate	0.2969	0.2257
Demographic Parity Ratio	0.8261	
Class 4 Selection Rate	0.5300	0.6415

Table 5: Overall Metrics for Class 4 by Gender

4.2 Fairness

While analyzing feature distributions, we observed a contradiction between education level and gender. We further examined how fairness metrics and selection rates vary between the genders, focusing also on the 4 poverty classes. We designated the gender of the household head as our sensitive attribute for assessing fairness metrics. (refer to 2 tables above)

4.2.1 Demographic Parity & Selection Rate

We defined class 4 as the positive class and the others as negative for a clear interpretation of model performance. The model has a **demographic parity ratio of 0.8261**, indicating relatively fair performance in predicting non-vulnerable households across genders. However, females show a lower selection rate compared to males, suggesting potential pre-existing bias.

4.2.2 False Negative Rate & False Positive Rate

Classes 1 to 3 exhibit low precision and recall, highlighting difficulties in accurately classifying vulnerable groups. Males have higher precision and recall across most classes, except in class 1 where females show notably higher precision.

For Males: precision for classes 1 to 3 ranges from 0.09 to 0.37, indicating poor performance in accurately predicting vulnerable males. Their recall for these classes is also low (0.06 - 0.44), underscoring difficulties in correctly identifying vulnerable households.

For Females: precision and recall rates are slightly better than males, suggesting the model is more effective at predicting vulnerability in households headed by females. This is advantageous given that females are more likely to be in vulnerable households in the dataset. FPR is lower for females in class 1 but higher in class 4 compared to males, indicating fewer false identifications of vulnerability among females. However, the FNR is significantly higher for females in class 1 (0.9412) than males (0.5435), indicating frequent misclassification of vulnerable females. In contrast, a lower FNR for females in class 4 suggests better identifying non-vulnerable cases.

4.3 Feature Importance

We examined the feature importances as ranked by the LGBM model to understand how model interpretability aligns with performance, and to check if the feature rankings are appropriate.

4.3.1 Feature Importance Generated by LGBM

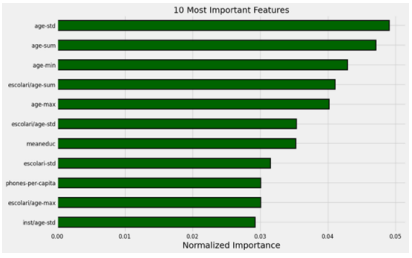

Figure 12: 10 Most Important Features

The author analyzed feature importances in the LGBM model, highlighting that age-related variables significantly influence the mode's poverty level predictions. Education variables, particularly **meaneduc** and **escolari/age-sum** (the ratio of years of schooling to age), also rank

highly, emphasizing the model's weighting on education and age. We have also listed out the ranking of variables in ascending order in our Colab. Dependency is another key variable impacting predictions. While the relevance of education and dependency in determine poverty levels is logical (higher education correlates with lower poverty, higher dependency with higher poverty), the prominence of age as a predictor becomes less crucial, especially when other critical factors like house condition and number of children should also be considered. This suggests that while the LGBM model performs well overall, it may not effectively address the nuanced needs of households in poverty.

4.3.2 SHAP

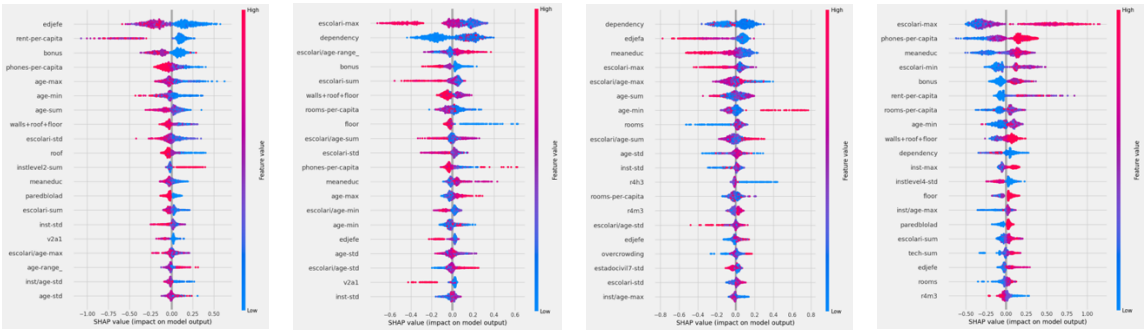

Figure 13-16: Summary Plot for Class 1-4

- **Class 1 (Extremely Vulnerable):** Primary features include **edjefe** (years of education for male head) where less education correlates with higher vulnerability. The average rent per person (**rent-per-capita**) suggests that household affording higher rent are less likely to be extremely poor. Additionally, **bonus**, representing amenities like fridges and tablets, are less common in these households, indicating higher vulnerability.
- **Class 2 (Moderate Poverty): dependency** and **escolari-max** (max years of schooling) are critical, with more schooling reducing likelihood of this classification. Poor house condition also increases the probability of being classified as moderate poverty.
- **Class 3 (Vulnerable): dependency** is a significant but less impactful feature, with unclear effects. **edjefa** (years of education for female heads) and other education-related features show that higher education levels in female heads reduce vulnerability. Additionally, the presence of more females in the household **r4m3** increases the likelihood of being classified as vulnerable.
- **Class 4 (Non-Vulnerable):** Higher education levels and rent payments correlate with a lower probability of vulnerability; Higher quality house led to higher probability of being classified as non-vulnerable

SHAP values offer detailed insights into feature importances, highlighting the model's stronger performance in predicting non-vulnerable households compared to those in poverty. Education heavily influences the model's outcomes, and house conditions are key indicators of poverty levels. Notably, the model tends to classify households with more females as vulnerable (class 3), suggesting pre-existing and technical bias that emphasizes gender in poverty predictions.

5. Summary

After evaluating the model's accuracy and fairness, we believe that the ADS is neither accurately nor fairly predicts vulnerable households. The intervention of the ADS should be assisting vulnerable households in need; however, the model performs poorly overall in predicting vulnerable households. This outcome might benefit stakeholders like the local government or public agencies by reducing expenditures on social welfare programs. However, the model adversely affects vulnerable household by failing to provide them necessary assistance. Metrics such as precision, recall, FNR, and FPR provides insights into how effectively they assist households. Additionally, the demographic parity ratio helps ensure that assistance is equitably distributed across genders among vulnerable households.

Although the model's performance appears similar across genders, key disparities should be noted. The number of females in a household negatively impacts poverty predictions, and the years of education for males and females affect poverty levels differently. This not only highlights issues in poverty prediction but also underscores existing gender inequities in society. These findings suggest a pre-existing bias where female may be treated unfairly despite having higher level of education, and household with more females are more vulnerable.

A major challenge in this competition is the lack of data for vulnerable households. The highly imbalanced distribution leads to model to overfit to non-vulnerable classification and fails the purpose of accurately predicting households in need. Improvement could include collecting more robust data that better reflects individual and household characteristics as well as vulnerable households. Implementing oversampling of minority classes during model development could help mitigate this imbalance and reduce overfitting.

Applying this model in public sector could be detrimental as it fails to effectively predict and assist households in need. Stakeholders like vulnerable Costa Rican households will be negatively impacted as their needs may be overlooked due to the model's shortcoming. Furthermore, data collection for vulnerable households and proxy means test are already challenging tasks. Using incomplete data with current method will lead to emergent bias where inaccurate predictions are produced as result of using imbalanced data. Public sectors could try to collect more relevant features that are more directly related to poverty levels.

6. Work Cited

- **Kaggle Competition:** "Costa Rican Household Poverty Level Prediction." Kaggle, www.kaggle.com/competitions/costa-rican-household-poverty-prediction/overview. Accessed 9 May 2024.
- **Link to Google Colab Notebook:** https://tinyurl.com/pb92xx52
- **Kaggle Solution:** Koehrsen, Will. "A Complete Introduction and Walkthrough." Kaggle, Kaggle, 19 Aug. 2018, www.kaggle.com/code/willkoehrsen/a-complete-introduction-and-walkthrough#Costa-Rican-Household-Poverty-Level-Prediction. Accessed 9 May 2024.

Appendix A: Partner Project Contributions

Kelly and Yui shared equal responsibility and contributed collaboratively to all the project:

- **ADS Selection and Proposals:** Joint effort by Kelly and Yui
- **Reproducing Original Solution:** Kelly handled preprocessing, Yui replicated the model
- **Accuracy Analysis:** Led by Kelly
- **Fairness Analysis:** Led by Yui
- **Robustness, Summary, and Findings:** Jointly done by Kelly and Yui
- **Report:** Jointly contributed by Kelly and Yui
- **Presentation and Slides:** Jointly contributed by Kelly and Yui

Note for graders: We noticed the professor's submission edit after consolidating our report to 10 pages. Although partner project contributions are not the core focus of the report, we've included it as appendix here (hope it won't breach the 10-page limit). Thank you so much for the amazing semester! We are really grateful for your time and help throughout the months, especially with office hours/labs/homeworks/exams etc. Both of us have learned so many valuable insights about RDS and this project on auditing ADS really provides us with a hands-on opportunity to apply our knowledge from the class effectively and we enjoyed it a lot. Hope you have a great summer!