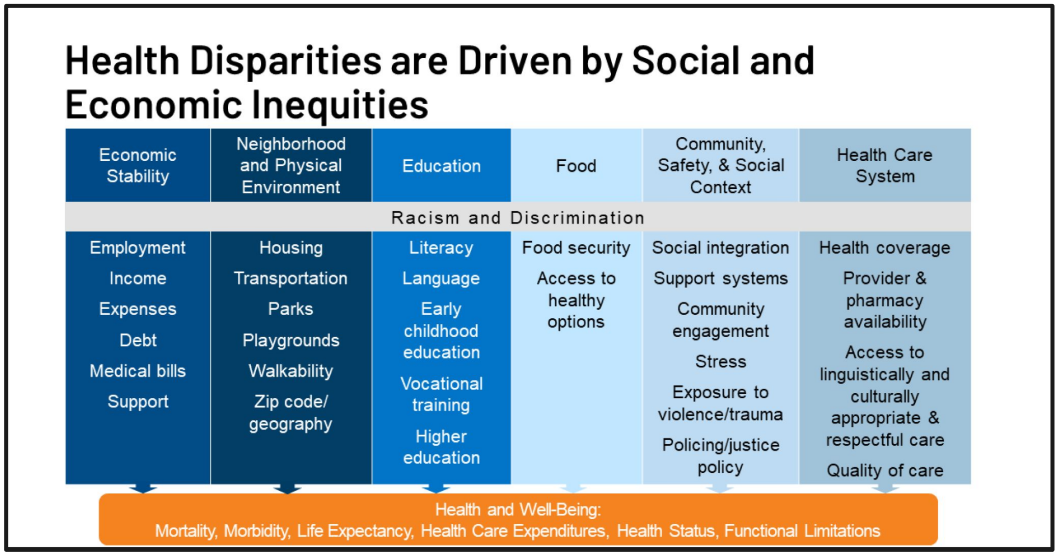
Machine Learning for HIV-ART Optimization

Context

- HIV (human immunodeficiency virus): A disease that reduces an individual's immune system and destroys cells used for battling infections.
- There is no cure to HIV but there are various treatment plans that help to prevent it.
- HIV can cause AIDS. This is the most common in children.
- According to hiv.gov, approximately 1.5 million children under the age of 15 had HIV in 2022.
- There is a treatment plan called ART: antiretroviral therapy
  This treatment plan suppresses the HIV from replicating.
- With proper care and use, this plan will eventually reduce mortality rates

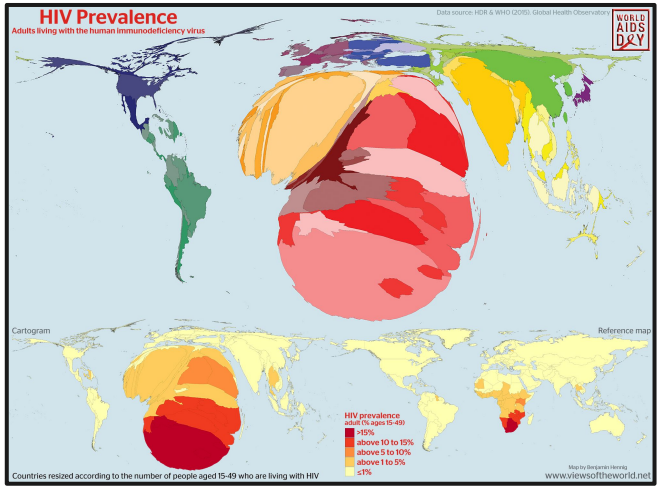
Problem Statement

- There exists a notable discrepancy in healthcare provision between developed Nations and those that are underdeveloped



Purpose

- The purpose of this project is to identify areas that are in need of ART treatment for children.
- I used AI machine learning models in order to predict which countries are in need of ART treatment and how much that country influences HIV on a global scale.



Data

- The data we used is called "hiv art coverage"
- This data consisted of the data from 2010 to 2022.
- It consists of the regions, the global influence and the ART coverage based on the region.

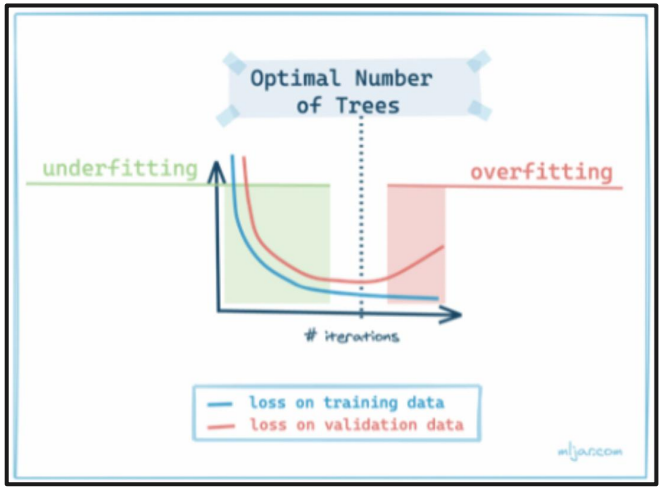
Country_Region	Year	ART_Coverage
135 unique values	2020.80 - 2022.00 Count: 457 2010 2022	0 990
Global	2010	16.7
Global	2011	21.1
Global	2012	25.4
Global	2013	29.4

Model

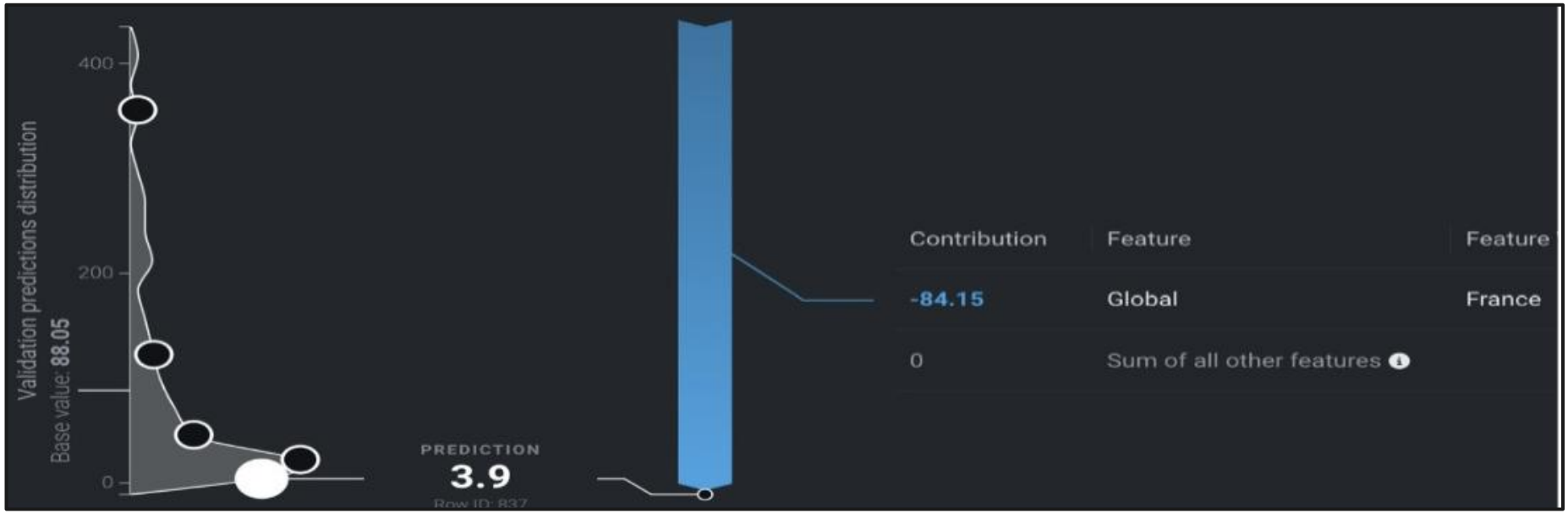
- I used a platform called DataRobot
- DataRobot allows us to upload various sets of data in order to implement features that allowed me to train the model to get accurate results.
- We used various features such as:
  -Regression model
  -A Quick Autopilot modeling mode
  -RMSE optimization metric
  -User-defined grouping
  -Cross Validation

Model Results

- Based on the features and filters, we used a **Light Gradient Boosted Trees Regressor with Early Stopping** and **Generative Additive2 Model**
- This model is an ensemble technique that combines multiple weak learners, typically decision trees, to create a robust and powerful predictive model. It does so in an iterative fashion, where each new stage (tree) corrects the errors of the previous ones.

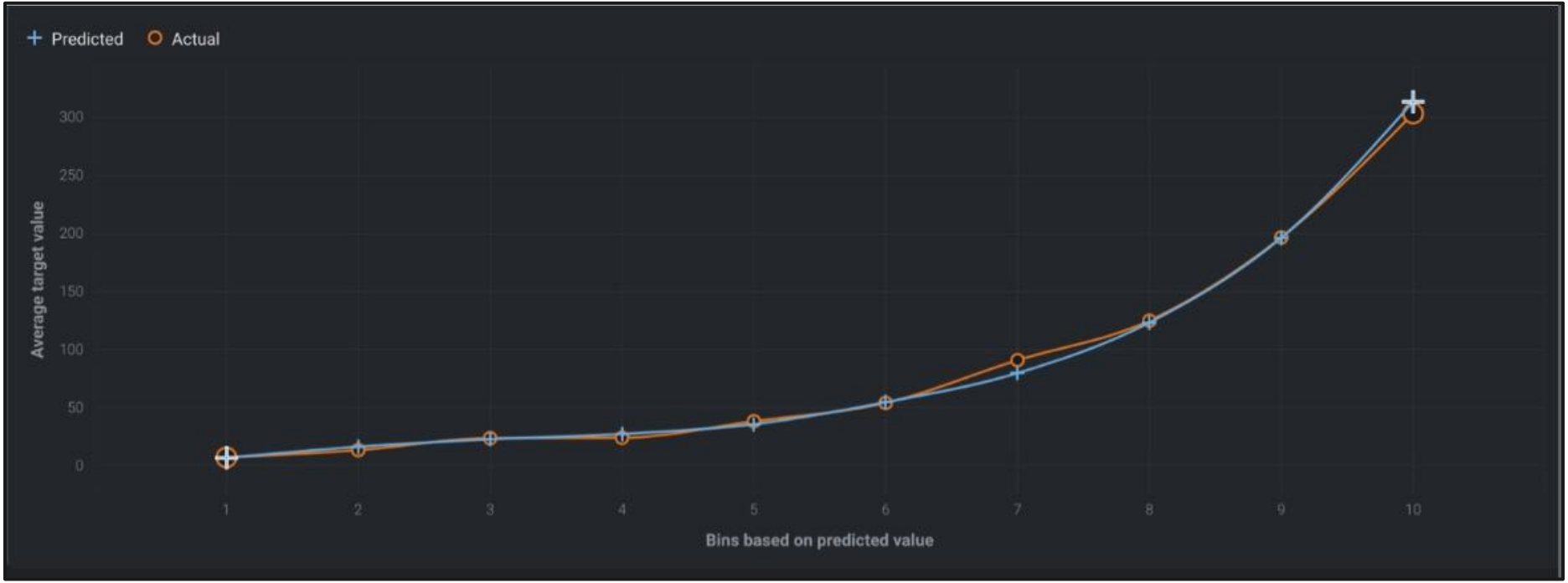


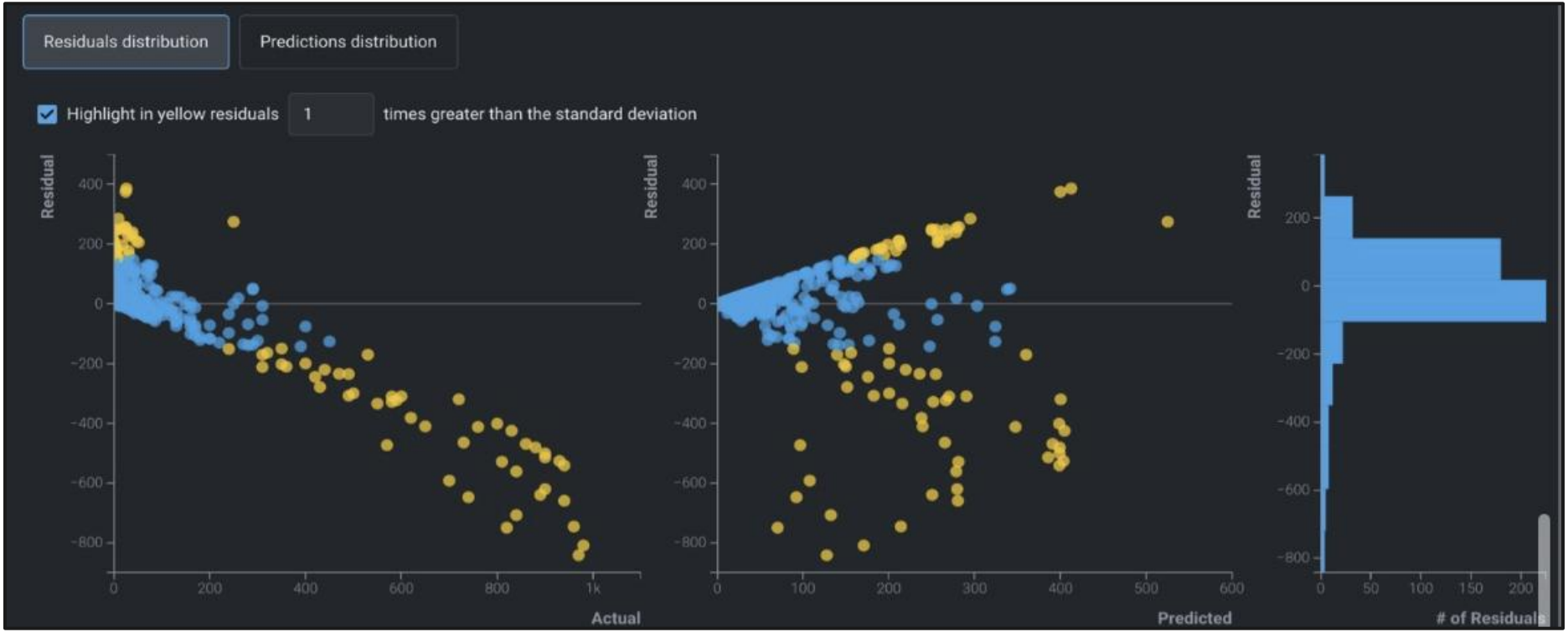

400
200
0
Validation predictions distribution
Base value: 88.05
PREDICTION
355.68
Row ID: 1044
Contribution
Feature
Feature Value
267.63
Global
Guatemala
0
Sum of all other features


Validation predictions distribution
Base value: 88.05
400
200
0
PREDICTION
3.9
Row ID: 837
Contribution
Feature
Feature
-84.15
Global
France
0
Sum of all other features

Lift Chart:

The Lift Chart Depicts how well a model segments the target population and how capable it is of predicting the target, letting you visualize the model's effectiveness.




Residuals distribution
Predictions distribution
Highlight in yellow residuals
1
times greater than the standard deviation
Residual
400
200
0
-200
-400
-600
-800
0
200
400
600
800
1k
Actual
Residual
400
200
0
-200
-400
-600
-800
0
100
200
300
400
500
600
Predicted
Residual
200
0
-200
-400
-600
-800
0
50
100
150
200
# of Residuals

Application

- By being addressed through the results of the machine learning model, governments can make efficient decisions for resource allocation towards specific countries
- The data can aid in providing early action government treatment for children as classified by HIV for efficient healthcare
- The model can constantly be improved to fit relevancy and allow for collaboration with other organizations in a country
- The models allow for informed decisions and the impact of that bring about effective health campaigns and where governments should intervene

Bibliography

- https://www.google.com/url?sa=i&url=https%3A%2F%2Fmljar.com%2Fblog%2Fxgboost-early-stopping%2F&psig=AOvVaw1HgwO2jJRsPR6yFMVM0D9D&ust=1708214474676000&source=images&cd=vfe&opi=89978449&ved=0CBMQjRxqFwoTCLCYhvqlsYQDFQAAAAADAAAAABAD
- https://www.google.com/url?sa=i&url=https%3A%2F%2Fourworldindata.org%2Fgrapher%2Fhiv-deaths-averted-from-antiretroviral-therapy-art&psig=AOvVaw1Z67YWDxcq246pgrzduUkJ&ust=1708214404053000&source=images&cd=vfe&opi=89978449&ved=0CBMQjRxqFwoTCNDOsruOsYQDFQAAAAADAAAAABAD
- https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.viewsoftheworld.net%2F%3Fp%3D5015&psig=AOvVaw35AIGu3quXPgXtrpq6Vy_v&ust=1708214300142000&source=images&cd=vfe&opi=89978449&ved=0CBMQjRxqFwoTCMD5v8WOSYQDFQAAAAAAAAAABAD

- "Hiv around the World - Google Search." Google.com, 2019, www.google.com/search?sca_esv=c33fa338858c7c86&q=hiv+around+the+world&tbm=isch&source=lnms&sa=X&ved=2ahUKEwjC3-SjiLGEAXWD120FHXaOCOYQ0pQJegQICxAB#imgrc=exSODpiB47aPLM. Accessed 16 Feb. 2024.
- "Light Gradient Boosted Trees Regressor with Early Stopping - Google Search." Google.com, 2019, www.google.com/search?q light+gradient-boosted+trees+regressor+with+early+stopping+&tbm=isch&ved 2ahUKEwjdsqbXiLGEAxWrzekDHbb2BIkQ2-cCcgQIABAA&oq=light+gradient+boosted+trees-regressor+with+carly+stopping+&gs_lp=EgNpbWciO2xpZ2h0IGdyYWRpZW50IGJvb3N0ZWQgdHJ1ZXMgcmVncmVzc29yIHdpdGggZWFybHkgc3RvcHBpbmcgSPZmUJgPWIFjcAF4AJABAZgBzwKgAZE3qgEJMTQuNDUuMS4xuAEDYAEA-AEBigILZ3dzLXdpeil pbWfCAgUQABIABMICCBAAGIAEGLEDwgIKEAAYgAQYigUYQ8ICDRAAGIAEGIOFGEMYSQPCAgeQABIABBgYiAYB&sclient=img&ei=BPfPZd2gA6ubp84Ptu2TyAg#imgrc 16Xs_aDqV0Ot8M. Accessed 17 Feb. 2024.
- "Number of Hiv Deaths Averted from Antiretroviral Therapy (ART), 2022 - Google Search." Google.com, 2022, www.google.com/search?q=number+of+hiv+deaths+averted+from+antiretroviral therapy (ART) %2C+2022&tbm=isch&ved=2ahUKEwi21OCliLGEAxUwHNAFHS9VB3gQ2-cCegQIABAA&oq=number+of+hiv+deaths averted+from+antiretroviral therapy(ART)%2C+2022&gs_lp=EgNpbWciQ25 lbWJlciBvZiBoaXYgZGVhdGhzIGF2ZXJ0ZWQgZnJvbSBhbnRpcmV0cm92aXJhbCB0aGVyYXB5KEFSVCksIDIwMjJI384BUMcHWMXGAXALACQAQGYAfwDoAG4TaoBCzEwLjYzLjMuNS0xuAEDYAEA-AEBigILZ3dzLXdpeil pbWICAgoQABIABBIKBRhDwgIGEAAYCBgewgIHEAAYgAQYGMICCBAAGIAEGLEDwgIFEAAYgATCAg0QABIABBIKBRIDGLEDIAYB &sclient=img&ei=nPbPZbbWCLC4wN4Pr6qdwAc#imgrc=N-6uOHRPI7ZnLM. Accessed 17 Feb. 2024.