Beyond poverty as a proxy: reducing inequality in infant mortality by identifying and targeting higher risk births

Antonio P. Ramos* Chad Hazlett† Stephen Smith‡

July 20, 2021

Abstract

Infant mortality remains high and uneven in much of sub-Saharan Africa. Given finite resources, reducing premature mortality requires effective tools to identifying left-behind populations at greatest risk. While countries routinely use income- or poverty-based thresholds to target policies, we examine whether models that consider other factors can substantially improve our ability to target policies to higher-risk births. Using machine learning methods, and 25 commonly available variables that can be observed prior to birth, we construct child-level risk scores for births in 22 sub-Saharan African countries. We find that targeting based on poverty, proxied by income, is only slightly better than random targeting, with the poorest 10 percent of the population experiencing approximately 10 percent of total infant mortality burden. By contrast the 10 percent of the population at highest risk according to our model accounts for 15-30% of infants deaths, depending on country. A hypothetical intervention that can be administered to 10% of the population and prevents just 5% of the deaths that would otherwise occur, for example, would save roughly 841,000 lives if targeted to the poorest decile, but over 1.6 million if targeted using our approach.

*Research Scientist, Department of Biostatistics, Fielding School of Public Health, California Center for Population Research University of California, Los Angeles. Email: tomramos@ucla.edu(corresponding) URL: http://www.antoniopedroramos.com/

†Assistant Professor, Departments of Statistics and Political Science, University of California Los Angeles. chazlett@ucla.edu URL: http://www.chadhazlett.com.

‡PhD Candidate, Department of Statistics, University of California Los Angeles. Email: stephensmith13424@gmail.com.

1 Introduction

Goal 3.2 of the Sustainable Development Goals (SDG) seeks to "end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to at least as low as 12 per 1,000 live births and under-5 mortality to at least as low as 25 per 1,000 live births." Despite large overall improvements, including a 44% reduction in child mortality globally from 2000 to 2015, progress has varied widely both between and within countries country to country (Rajaratnam et al., 2010; Stuckler et al., 2010; Deaton, 2020), with under-5 mortality exceeding 80 per 1,000 live births in some countries of sub-Saharan Africa. Even improvements in national averages are often accompanied by widening gaps among subpopulations, with more privileged groups improving their health status faster than others (Bendavid, 2014; Moser et al., 2005)

Many of these deaths are preventable with existing, low-cost technologies and interventions (Black et al., 2003; Jones et al., 2003). However, employing these solutions under budget constraints requires targeting them to those who would have otherwise died. The statistical rareness of early-life mortality thus severely limits the effectiveness of any intervention. Although multiple risk factors play a role in early-life mortality (see e.g. Mosley and Chen, 1984), many countries target policies and interventions based on one or a few risk factors, almost invariably indicators for low income or poverty. For example, cash transfer programs designed to reduce early mortality are often targeted to the poor (e.g. Glassman et al., 2013; Basset, 2008; Akresh et al., 2015). In Burkina Faso, families enrolled in conditional cash transfer schemes were required to obtain quarterly child growth monitoring at local health clinics for all children under 60 months of age (Akresh et al., 2015). The randomized controlled trial (RCT) *lentils for vaccines* in India targeted the poor, as have most RCTs that aim to increase vaccine uptake, good nutrition, or child health more generally (Banerjee et al., 2010). Yet, such income- or poverty-based eligibility requirements may do

little to target those actually at highest risk. For example, Ramos et al. (2020) finds in Brazil and India that 65% of all child deaths are not not among the 20% poorest, with higher values in many other countries. Such poor targeting severely limits the potential effectiveness of any life-saving intervention that can be distributed only to those considered higher risk by these approaches, failing to save as many lives as it could or to reduce inequalities in mortality risk.

In this paper we employ a variety of machine learning methods and a wide range of risk factors to more accurately estimate mortality risk in infants (under one year of age) at the individual level in 22 countries in sub-Saharan Africa. Very few studies to date have employed more than a few variables in generating risk predictions of this kind (Ramos et al., 2018, 2020; Houweling et al., 2019). We employ only observable risk factors that we deemed would reasonably be available to policymakers or health-workers in practice. We also restricted the models to employ only information that could be used for planning purposes prior to a child's birth, i.e. excluding individual health data that would need to be collected post-birth. These approaches allow us to flexibly determine which risk factors to include in different countries based on their predictive power and data availability, while also flexibly modeling the relationship between these factors and mortality. We compare our results with models that are based only on income.

2 Methods

2.1 Data Sources

We use the most recent Demographic Health Survey (DHS) for each of 22 sub-Saharan countries, provided by IPUMS Global Health. Table 1 lists these countries and their number of observations. The data used in this study come from the most recent available Demographic and Health Surveys (DHS) (https://dhsprogram.com/). These are nationally representative

surveys that have been conducted in more than 100 low and middle income countries since 1988 and they are one of the most widely source of demographic and health data available for the poorest nations.

2.2 Data

Our outcome is infant mortality, indicated by death before reaching one year of age. For risk factors, we chose variables on two grounds: they must be reasonably available to policymakers and health workers in-country, and relatedly, should be measurable prior to the birth of a given child, rather than depending on the health or other features of the child once borne. The latter constraint is in place so that these risk factors could be used to calculate risk in advance and deploy interventions to households, clinics, or regions with high risk. The resulting risk factors we consider are maternal age, malaria prevalence in the vicinity, head of household age and sex, parental education, cooking fuel type, floor type, toilet type, drinking water type, income, birth order, birth month, religion, bed net usage, age of first marriage, and the death of a previous child under one or five years of age. Geographic information is also used, but varies in its granularity by country. We recode variables with many categories to construct measures that are likely to perform better. For example, the DHS data initially contain 66 categories of toilet types. We simplified this to 4 categories based on the level of improvement. Appendix A provides a complete description of these coding decisions. Table 1 indicates, for each country, the set of variables (if any) that is not available in the DHS data, the number of observations (original and with complete data), the number of infant deaths, and a description of the type of geographic data available.

Finally, the DHS does not contain an indication of how an individual or household is classified relative to the poverty line in a given country, nor does it attempt to measure income. Thus rather than including such a locally-defined poverty measure, we employ the wealth-based indices provided by DHS. This measure is a factor imputed by DHS, using

Table 1: Data availability by country

Country	N original	N complete	N deaths	Geography	Unavailable
Angola	11024	10829	472	province (18)	age first marriage, malaria
Benin	10490	10199	471	department (12)	
Burkina Faso	11728	10905	752	region (13)	
Cameroon	9041	8450	564	region or city (12)	
Congo (D.R.)	14384	12731	785	province (11)	previous child death(s)
Cote d'Ivoire	5945	5211	421	none	
Ghana	4548	4203	189	none	
Guinea	5422	5223	396	region (8)	cooking fuel
Kenya	16557	15110	589	none	pregtermin
Lesotho	2360	2127	145	district (10)	bednets, malaria
Madagascar	9657	9229	459	region (22)	cooking fuel
Malawi	13633	13231	563	region (3)	malaria
Mali	8043	7913	475	region (6)	cooking fuel, malaria
Mozambique	8493	8026	537	province (11)	
Niger	9650	9612	586	region (8)	religion, malaria
Nigeria	24222	23559	1686	geopolitical zone (6)	
Rwanda	6055	5350	168	province (5)	cooking fuel
Senegal	9585	9371	401	region (14)	malaria
Tanzania	8007	7688	319	district (10)	cooking fuel, religion, malaria
Uganda	12089	11619	484	district (112)	cooking fuel, malaria
Zambia	10646	9811	483	none	
Zimbabwe	4852	4630	236	province (10)	malaria

Note: Summary of data by country. The DHS data initially include `N original` observations in each country, which falls to `N complete` when we remove missing values on the included variables. `Geography` indicates what type of geographic information, if any, is available and the number of units in parentheses. All other variables described in text are used in each country except where otherwise indicated under `Unavailable`.

principal components analysis over a number of variables including ownership of radios, televisions, and other domestic equipment; electricity and clean water; type of materials used in the walls, floor and roof; and the type of toilet in the household. Scores are calculated at the household level, separately by country.

2.3 Modeling

We estimate infant mortality risk for each birth in the data set. In each country, we separately run two sets of models. One is a "wealth model", described below as a proxy for what can be achieved through poverty-based targeting, and the other are flexible machine learn-

ing approaches that employ 22 variables to estimate individual level mortality risk. These include random forest (rf), extreme gradient boosting over trees (xgb), kernel regularized least squares (krls), and elastic-net regularized logistic regression (elastic-logit).

Both rf (Breiman, 2001) and xgb (Chen and Guestrin, 2016) are tree-based approaches. The simplest tree-based approaches are classification and regression trees (cart, Breiman et al., 1984), which partition the input variables recursively into groups that predict the outcome level. While single-tree models like cart are easy to interpret, they often perform poorly, and small changes in the input data can have dramatic effects on the model and results. The rf and xgb approaches both improve upon such single-tree approaches by combining many trees. For rf, this is achieved by effectively averaging together the estimates of many trees, each employing only a random subset of the input variables and trained on a random subset of the observations. As applied here, each forest contains 500 trees, and the number of variables randomly sampled for each tree split (`mtry`) was set for each country's model using cross-validation. Under xgb, this is similarly achieved by combining many trees, but by beginning with one tree, adding the predictions of another tree trained on the errors of the prior, then repeating this process hundreds or thousands of times. We use cross-validation for each country to determine (i) the learning rate ($\eta \in (0, 1)$, where smaller values help to prevent overfitting), (ii) the maximum depth of the trees (`max_depth`), (iii) the proportion of the variables to be considered for tree construction (`colsample_bytree`), (iv) the proportion of observations from the training set used for modeling (`subsample`), and (v) the number of iterations for the boosting procedure (`nrounds`). The remaining parameters were set at common default values: the minimum requirement of prediction improvement before selecting a more complex model (γ) was set to 0, and `min_child_weight`, where higher values restrict the depth of each tree based on a measure of the homogeneity of the labels within the nodes (a sign of overfitting), was set to 1.

Kernel-based regularized least squares (krls, Hainmueller and Hazlett, 2014) is a regres-

sion and classification technique that employs kernels. Such approaches effectively build a model by leveraging information about how similar observations are to each other. Each observation is treated as a potential "fence post" in the data, and the model learns how being closer to or further from each such location in the data influences the expected "height" (probability of mortality). We employ a Gaussian kernel as the measure of similarity using default values suggested in Hainmueller and Hazlett (2014). This approach also allows for a degree of interpretation, because it allows marginal effects to be determined similar to those reported in conventional regression tables.

Elastic-net logistic regression (elastic-logit, Zou and Hastie, 2005) is similar to conventional logistic regression. While ordinary logistic regression chooses coefficients β_j on each variable j so as to maximize the likelihood by minimizing the negative log-likelihood, elastic-net logistic regression minimizes the negative log-likelihood *plus* a penalty term that helps to regularize or "shrink" the model, preventing overfitting. In our case that penalty term is equal to $\lambda(\frac{\alpha}{2}\sum_j |\beta_j| + \frac{1-\alpha}{2}\sum_j \beta_j^2)$. The parameters α and λ are chosen by cross-validation. This constitutes a compromise between what is known as ridge regression (which is penalized by the ℓ_2 norm, $\sum_j \beta_j^2$), and LASSO (penalized by the ℓ_1 norm, $\sum_j |\beta_j|$). By including the ℓ_1 norm in the penalty, elastic-net has the benefit of setting some coefficients to exactly 0, that is, selecting some variables while dropping others.

Finally, we construct a simple ensemble model in which each probability is a simple unweighted average of the probabilities produced by elastic-logit, rf, xgb, and krls. This has two primary benefits. The first is that such ensembles often have performance as good as the best sub-models and sometimes better. Second, it allows us to commit to reporting results for a single model—the ensemble model—as our premier estimate of the predictive power of these model on the test data. This is preferable to the tendency to choose the best-performing model on the test set, which can be misleading because such a choice is effectively a means of "training on the test set".

Wealth model

Our wealth-based model employs wealth information flexibly: mortality is regressed on ten indicators for decile of wealth, equivalent to finding the mean mortality level in each wealth decide. In evaluating model performance, we can consider how being in the highest risk decile, quartile, or other groups is related to mortality risk. There is no presumption that risk is necessarily decreasing or even monotonic in the level of wealth.

We note that this is a wealth model, not a poverty model based on poverty assessments made by countries, usually based on income. Neither income nor poverty-assessments themselves are available in the DHS data. This limits our ability to benchmark against a true poverty-based system. That said, our wealth model may reasonably provide an optimistic indication of what can be predicted by income or poverty-based approaches. Household income is time-sensitive, but is likely to generate risks of early mortality when it remains low on average over a long period of time, in which case wealth will also drop. Moreover, financial shocks such as large expenses or losses may have little impact on income, but substantial impact on quality of life, access to healthcare, and poverty as locally understood (see e.g. Krishna, 2004). Our wealth model is also more flexible and less coarse than poverty-based targeting systems, which typically assign individuals through a binary system (e.g. above or below the poverty line).

Sample splitting and cross-validation

Before fitting each model, we split the data into training and test sets. For each country, 80% of observations are placed in a training and validation set, while 20% are retained in a held-out testing set. For the models described above, we tune hyperparameters using either 10-fold cross-validation within the training set (for elastic-logit, rf, xgb) or leave-one-out cross-validation (for krls). The cross-validation procedure also gives a preliminary assessment of model performance. Once all modeling choices were finalized based on this

cross-validation, those models were "frozen" and we then assessed performance on the test set. This was done only once, after writing the bulk of the paper based on the cross-validation results, then updating the final values using the actual test-set results. This provides an honest assessment out-of-sample performance that prohibits "training on the test set". We report the accuracy of the models using the performance metrics described below.

2.4 Performance metrics

To evaluate our approach, we employ interpretable and policy-relevant metrics. A model's *recall* (also referred to as sensitivity) is the proportion of actual deaths that are included in the group we consider to be at risk. What we term "recall10" is defined by first determining the group of births in the top risk decile (hence the "10") according to our model, then computing the fraction of all actual infant deaths accounted for by this group. In policy terms, this corresponds to a case in which we have resources to target 10% of the population, and wish to know how much of the total mortality risk would be covered in that group. An ineffective risk assessment that assigns risk at random would produce a recall10 of 0.10 (10%). By contrast an effective system points us to a top risk decile that accounts for more than 10% of the mortality, perhaps 20% or 30%, for example. Similarly, we report recall20, which tells us the fraction of early deaths that occurred within the top risk quintile. For both recall10 and recall20, we also report "efficiency gain", which tells us how many times more effective the ensemble model is than the benchmark wealth model.

2.4.1 Variable importance

Our models only seek to determine the best risk estimate possible given a set of predictors and the answers say nothing of the causal effect of those predictors on the risk of mortality. Indeed many of the risk factors, such as religion or the type of roof, are valuable not because they cause mortality but because they are expected to signal the presence of other, unobserved

factors that cause higher mortality. In other words, confounding is expected and hoped for in these models.

It is nevertheless of value to know what the models are doing to generate forecasts given the input variables. In particular we may wish to know which variables are most important in estimating the risk, for the purposes of assessing the credibility of the model or in focusing public health resources on data collection. Unfortunately, for many machine learning models it is more difficult to characterize what the model's behavior than is the case with simpler regression approaches like OLS, in which the model is fully characterized by a small number of parameters. However, some tools are readily available and provide insights into the more and less important variables. Specifically, for elastic-net we report how often the model choose a given variable and the proportion of time their coefficient is positive. For random forest, the (scaled) "variable importance" describes what fraction of the time a given variable was chosen to be included in the classification trees over which each random forest model aggregates.

3 Results

The primary results are recall10 and recall20 rates computed on the held-out test set (Tables 2 and 3).[1]

3.1 Poor performance of the wealth model

Table 2 shows the proportion of mortality that occurred among the top 10% highest predicted risk individuals. In Rwanda and Uganda, wealth is a reasonable indicator of risk, and the 10% of the population with the higher wealth-based predicted risk account for 24% and 18%

[1] Appendix A.1 provides the same performance metrics, but using cross-validated results from the training set. Those results are extremely similar to the test set results reported here, indicating very little over-fitting in the training process.

Table 2: recall10 results, test set

Survey	wealth	elastic-logit	rf	xgb	krls	ensemble	efficiency gain
Angola 2015	0.11	0.23	0.23	0.23	0.22	0.27	2.50
Benin 2011	0.13	0.22	0.31	0.20	0.21	0.22	1.75
Burkina Faso 2010	0.12	0.15	0.21	0.17	0.15	0.19	1.61
Cameroon 2011	0.10	0.18	0.23	0.17	0.17	0.24	2.45
Congo Democratic Republic 2013-14	0.13	0.14	0.18	0.19	0.14	0.18	1.40
Cote d'Ivoire 2011	0.12	0.21	0.17	0.27	0.20	0.24	2.00
Ghana 2014	0.05	0.11	0.19	0.16	0.08	0.19	3.50
Guinea 2012	0.11	0.14	0.20	0.16	0.15	0.19	1.67
Kenya 2014	0.06	0.21	0.17	0.12	0.19	0.21	3.57
Lesotho 2014	0.14	0.14	0.14	0.14	0.10	0.17	1.25
Madagascar 2008	0.08	0.10	0.15	0.10	0.13	0.21	2.71
Malawi 2016	0.12	0.12	0.21	0.20	0.12	0.19	1.62
Mali 2012	0.07	0.20	0.21	0.23	0.23	0.23	3.14
Mozambique 2011	0.09	0.13	0.21	0.13	0.13	0.21	2.20
Niger 2012	0.12	0.15	0.13	0.17	0.13	0.16	1.36
Nigeria 2013	0.12	0.21	0.22	0.19	0.24	0.25	2.15
Rwanda 2014	0.24	0.21	0.09	0.21	0.21	0.21	0.88
Senegal 2017	0.10	0.16	0.26	0.29	0.15	0.29	2.87
Tanzania 2015	0.03	0.21	0.14	0.16	0.17	0.22	7.00
Uganda 2016	0.18	0.14	0.24	0.24	0.25	0.26	1.47
Zambia 2013	0.10	0.20	0.23	0.16	0.19	0.21	2.00
Zimbabwe 2015	0.13	0.13	0.21	0.19	0.17	0.19	1.50
Average	0.11	0.17	0.20	0.19	0.17	0.22	2.30

Table 3: Recall20 results, test set

Survey	wealth	elastic-logit	rf	xgb	krls	ensemble	efficiency gain
Angola 2015	0.27	0.45	0.34	0.38	0.43	0.44	1.64
Benin 2011	0.28	0.36	0.46	0.34	0.39	0.46	1.65
Burkina Faso 2010	0.26	0.35	0.26	0.33	0.41	0.36	1.38
Cameroon 2011	0.29	0.27	0.33	0.29	0.29	0.34	1.19
Congo Democratic Republic 2013-14	0.20	0.26	0.32	0.29	0.29	0.31	1.55
Cote d'Ivoire 2011	0.25	0.44	0.26	0.40	0.36	0.43	1.71
Ghana 2014	0.16	0.16	0.22	0.27	0.16	0.24	1.50
Guinea 2012	0.18	0.27	0.34	0.24	0.33	0.33	1.86
Kenya 2014	0.17	0.26	0.32	0.25	0.31	0.32	1.90
Lesotho 2014	0.24	0.28	0.28	0.24	0.14	0.24	1.00
Madagascar 2008	0.16	0.26	0.30	0.24	0.25	0.31	1.87
Malawi 2016	0.26	0.23	0.30	0.31	0.25	0.30	1.17
Mali 2012	0.17	0.38	0.35	0.36	0.37	0.39	2.31
Mozambique 2011	0.21	0.23	0.32	0.22	0.29	0.34	1.64
Niger 2012	0.21	0.27	0.21	0.29	0.24	0.28	1.38
Nigeria 2013	0.22	0.39	0.35	0.37	0.40	0.40	1.81
Rwanda 2014	0.48	0.30	0.33	0.39	0.36	0.33	0.69
Senegal 2017	0.19	0.30	0.38	0.39	0.30	0.40	2.13
Tanzania 2015	0.11	0.32	0.29	0.35	0.25	0.29	2.57
Uganda 2016	0.25	0.27	0.33	0.39	0.34	0.42	1.67
Zambia 2013	0.26	0.32	0.34	0.31	0.32	0.31	1.20
Zimbabwe 2015	0.19	0.26	0.26	0.32	0.21	0.36	1.89
Average	0.23	0.30	0.31	0.32	0.30	0.35	1.62

of the mortality, respectively. In the other 20 countries, wealth is almost uninformative, with Recall10 values falling between just 3% (Tanzania) and 14% (Lesotho). Overall, the (country-wise) average recall10 of just 11% suggests that the wealth generally model does

little or no better than chance in predicting which individuals are at highest risk.

Similarly, Table 3 shows that in most countries, recall20 is only slightly above 20%, averaging (country-wise) to 23%. Rwanda is again an exception, with the top quintile accounting for 48% of mortality, while recall20 rates range from 11% to 29% in the others, and are at or below chance (20%) in 9 of the 22 countries.

3.2 Beyond wealth: Gains from richer models

The elastic-net logit, rf, xgb, and krls all augment the wealth model by adding the additional variables described above. These models perform similarly to each other overall, and far better than the wealth-only model. The top risk deciles according to these models account for, on average, 17-20% of mortality (recall10); the top risk quintile accounts for 30-32% of mortality (recall20). The ensemble model performs better than any of the individual models, on average capturing 22% of the mortality in its top risk decile (recall10), and 35% in the top risk quintile (recall20). Compared to the wealth model, the ensemble model also captures 2.3 times as much of the mortality in its top risk decile (efficiency gain for recall10) and 1.6 times the mortality in its top risk quintile (efficiency gain for recall20). No single country always shows the highest recall10 across models, nor the lowest. Moreover, within given countries, the ensemble model outperforms the constituent models in many cases, does as well as the top one or two models in many others, and is never the worst performing model in that country. Thus, relying on the ensemble model not only yields better results than any other model on average, but also safeguards against cases where individual models perform poorly in a given country.

3.3 Variable selection

The elastic-net logistic regression model for each country automatically chooses a subset of variables to retain in the model, dropping the others. Surprisingly, every variable we provided was included in at least 36% of the country-level models. That is, at least taken across countries, there are not just a few key variables; rather the models tend to be fairly dense such that every variable is present in at least a third of the models. A few variables, though, are included especially often: most models included the previous death of a sibling in the first year (86%) or the death of any previous sibling at any point (76%). 77% of country level models included the child's gender; 73% included an urban/rural indicator. Over two-thirds (68%) of models included the mother's years of education, and many included the birth month (73% use sine-transformed birth month; 55% use cosine-transformed birth month). Maternal age was included in 64% of models.

For random forest, the "scaled variable importance" describes what fraction of the time a given variable was chosen to be included in the classification trees over which each random forest model aggregates. Compared to elastic-net logistic regression, random forest was somewhat more selective, with over half of variables appearing in fewer than 15% of the regression trees. A few variables were particularly rarely included, such as clean cooking fuel (3%) and minority religion (5%). By contrast, maternal age was included nearly every time (99.6%), followed closely by wealth percentile and it's square (93%, 94%). The latter is particularly notable, as wealth percentiles were among the *least* commonly employed in the elastic-net logit models. Other important variables included malaria incidence rate in the area (92%), age of household head (77%), and mother's age at first marriage or union (63%).

Table 4: Variable Importance

Variable	**Elastic survival**	**Proportion of positive coefficients**	**rf variable importance**
previous death of sibling in first year	0.86	1.00	0.11
mother's years of education	0.77	0.06	0.39
sine-transformed birth month	0.73	0.69	0.46
female	0.73	0.00	0.17
previous death of sibling	0.71	1.00	0.11
rural	0.64	0.71	0.10
malaria	0.62	0.75	0.92
maternal age	0.59	0.08	1.00
previous pregnancy termination for mother	0.57	0.75	0.12
clean water access	0.55	0.58	0.14
cosine-transformed birth month	0.55	0.33	0.48
log-transformed birth order	0.55	0.25	0.54
mother's age at first marriage/union	0.52	0.45	0.63
log-transformed number of bednets	0.52	0.09	0.38
safe floor material	0.50	0.18	0.11
age of household head	0.50	0.64	0.77
female head of household	0.50	0.55	0.11
minority religion	0.50	0.70	0.06
christian	0.35	0.43	0.08
muslim	0.35	0.71	0.06
unimproved toilet	0.32	0.86	0.09
clean cooking fuel	0.31	0.20	0.05
square of wealth percentile	0.27	0.00	0.94
pit toilet	0.23	0.60	0.10
wealth percentile	0.23	0.60	0.93

4 Discussion

Across 22 countries in sub-Saharan Africa, predicting early mortality based on wealth measures alone was ineffective: those in highest 10% and 20% risk groups accounted for only 11% and 23% of mortality respectively. Fortunately, however, when wealth information is combined with other risk factors, flexible models can more accurately assess individual-level

risk of early mortality. On average across countries, the 10% of births with the highest risk as predicted by the ensemble model account for 22% of deaths, and the 20% at highest risk account for 35%. Compared to the wealth-only models—which we expect are optimistic benchmarks for the poverty-based targeting approach—each country's ensemble model on average would have identified 2.3 times as many deaths in the top risk decile, and 1.6 times as many in the top 20% risk group. Senegal offers a relatively representative example, in which targeting the poorest 10% will identify only 10% of all infant deaths, while targeting the 10% at highest risk according to the ensemble model would identify 29% of all deaths.

The various models used to incorporate these risk factors—a regularized logistic regression (elastic-logit), random forest (rf), extreme gradient boosting (xgb), and kernel regularized least squares (krls)—have similar average performance. However, some models exceed others in particular countries. Using the average of predictions from these (ensemble) generates better performance on average while importantly protecting against cases where individual models performed poorly in particular countries. Particularly important to many models were variables regarding the previous death of siblings in the first year, mother's age and education, malaria prevalence, and the child's gender. The value of these variables as predictors says nothing of their causal influence, and we often expect they are predictive because they signal the presence or absence of other, hard to measure factors that influence risk.

Central to the policy-relevance of such an approach is the question of whether the variables needed to run these models are generally accessible. The risk factors we included were all chosen to be feasibly collected and employed in targeting strategies by health agencies. We avoided post-birth variables, such as birth-weight, which require real-time data. Only the sex of the child and the calendar month of birth are unknown until the birth occurs, and this would not impeded the planning or siting of interventions and materials that are expected to reduce preventable deaths.

These findings are consistent with previous studies that show that births at high risk

of death exist across all socioeconomic groups, and that combining multiple risk factors improves our ability to more accurately identify the higher risk births (Ramos et al., 2018, 2020; Houweling et al., 2019). Our findings are related to a large body of literature in medicine and public health that develops risk scores to identify those at risk of some event (e.g. Mpimbaza et al., 2015; Beymer et al., 2017). To our knowledge, prior work on infant mortality using flexible models or machine learning approaches had not previously been used to develop birth-level risk scores in sub-Saharan Africa.

How beneficial might this improvement in predictive power be? We consider the potential benefit of an improved targeting system such as this for a hypothetical intervention under resource constraints. Consider an intervention that a given country can afford to administer to only 10% of births. Ideally, this would be targeted to the 10% at highest risk of early mortality. This intervention would not reduce mortality among children who would not have died anyway, but let us suppose it reduces mortality by some proportion, efficacy, among those who would have otherwise have died. In a country with N_{births} per year, the number of lives saved per year would be efficacy $* \, Pr(\text{death}|\text{high risk}) * N_{births}/10$. Applying Bayes' rule, this is simply efficacy $*$ recall10 $* \, Pr(\text{death}) * N_{births}$ where $Pr(\text{death})$ is the overall mortality rate expected for children born in that year (absent the intervention). Notice that the number of lives saved is proportional to the recall10 of the model used. Doubling the recall10—which is roughly what we see for all of our models compared to the wealth model—would double the lives saved by that intervention.

Using recent estimates of the number of births and baseline mortality rates in each of these countries from World Bank Development Indicators (2021), the number of lives saved by an intervention with efficacy = 5% would be 841 thousand for the wealth model, but 1.61 million for the ensemble model. In simpler terms, the efficiency gain of roughly two for these models compared to the wealth model translates into roughly double the lives saved by an

intervention with any particular efficacy.[2]

We emphasize that all models here were trained and tuned by cross-validation on a training set, and the performance indicated here represents their performance on a test set that was completely unseen during training. If such models are to be used in practice, they would be retrained on the available data in a given country at present, but we emphasize the importance of careful tuning by cross-validation to avoid over-fitting that would lead to both over-confidence in the model's ability and, quite possibly, poor performance in reality. That the results of our cross-validated models on the training set is nearly identical to the results on the test set, however, suggests that it is possible to estimate how well a model is likely to perform on as-yet unseen data if one is careful during the cross-validation procedure.

4.1 Limitations and Opportunities

Our work is not without limitations and indeed it explores the limits of machine learning methods to predict infant death risk. The first limitation of note is simply that not all deaths can be predicted, particularly with readily available data. In general we find our models for the highest 10% risk group capture only 20-30% of mortality. This is beneficial and worth using for targeting since the effectiveness of an intervention is proportional to this recall rate, as shown above. However, it leaves a great deal of unpredicted mortality spread across the lower risk groups. Thus, when possible, programs must still be targeted to a much wider group in order to capture a large fraction of births that will die.

Second, the DHS sampling procedure attempts to produce a sample that is representative of the population. If these are in fact not as representative as hoped, then while the results still reflect the predictive power of these models in some population of each country, the results may not be representative of these countries as a whole.

[2]Note that the average efficiency gain (2.3) is not exactly the ratio of lives saved because the latter takes the relative birth rates and mortality in each country into account.

Third, the link between the improved predictive power of risk models and improved outcomes such as lives saved is complicated. While our simple hypothetical example above is intended only as an illustration, that exercise also reveals some of these complications. First, while the assumed the "efficacy" of a given hypothetical intervention is applied only to those individuals who would have otherwise have died, it is nevertheless possible that the even in this sense the true efficacy would not be constant across risk levels. For example, when children from wealthier families die young, it may less often be due to preventable causes, making the efficacy of any likely intervention lower. In our case this does little to jeopardize our finding: the efficacy rate need only apply to those in the top risk decile, and doing a poorer job of estimating risk would only further reduce the lives saved by reducing the effective efficacy. That is, accounting for this would only amplify the comparative benefits we describe of improved targeting. Second and more complicated, we are not incorporating targeting costs in our analysis. The calculus of the efficiency gains assumes that interventions have the same costs for each birth. In reality, costs need to be adjusted according to local conditions. For example, geographic location could be a major factor in calculating costs. Targeting children clustered by geographic location is likely much cheaper and easier than targeting children distant from each other, for example.

Finally, in the future this approach could be further improved if information of the cause of death were available. Associating each birth to a particular cause of the death could generate models that suggest targeting births with more specific, needed interventions.

References

Akresh, R., de Walque, D., and Kazianga, H. (2015). Alternative cash transfer delivery mechanism: Impacts on routine preventive health clinic visits in burkina faso. working paper series 17785, National Bureau of Economic Research.

Banerjee, A., Duflo, E., Glennerster, R., and Kothari, D. (2010). Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *British Medical Journal*, 340(may17 1):c2220.

Basset, L. (2008). Can conditional cash transfer programs play a greater role in reducing child undernutrition? Technical report, World Bank.

Bendavid, E. (2014). Changes in child mortality over time across the wealth gradient in less developing counntries. *Pediatrics*, 134:e1551–e1559.

Beymer, M., Weiss, R. E., Sugar, C., Bourque, L. B., Gee, G. C., Morisky, D., Shu, S. B., Javanbakht, M., and Bolan, R. K. (2017). Are centers for disease control and prevention guidelines for preexposure prophylaxis specific enough? formulation of a personalized hiv risk score for pre-exposure prophylaxis initiation. *Sexually Transmitted Diseases*, 44(1).

Black, R., Morris, S., Bryce, J., and Venis, S. (2003). Where and why are 10 million children dying every year? commentary. *The Lancet*, 361:2226–34.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA.

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA. ACM.

Deaton, A. (2020). *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton University Press.

Glassman, A., Duran, D., Fleisher, L., Singer, D., Sturke, R., Angeles, G., Charles, J., Emrey, B., Gleason, J., Mwebsa, W., Saldana, K., Yarrow, K., and Koblinsky, M. (2013).

Impact of conditional cash transfers on maternal and newborn health. *Journal of Health, Population and Nutrition*, 31.4 Suppl 2 (2013):S48–S66.

Hainmueller, J. and Hazlett, C. (2014). Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach. *Political Analysis*, pages 143–168.

Houweling, T. A. J., Klaveren, D. v., Das, S., Azad, K., Tripathy, P., Manandhar, D., Neuman, M., Jonge, E. d., Been, J. V., Steyerberg, E., and Costello, A. (2019). A prediction model for neonatal mortality in low- and middle-income countries: an analysis of data from population surveillance sites in india, nepal and bangladesh. *International Journal of Epidemiology*, 48(1):86–198.

Jones, G., Steketee, R., Black, R., Bhutta, Z., and Morris, S. (2003). How many child deaths can we prevent this year? *The Lancet*, 362:65–71.

Krishna, A. (2004). Escaping poverty and becoming poor: who gains, who loses, and why? *World development*, 32(1):121–136.

Moser, K., Leon, D., and Gwatkin, D. (2005). How does progress toward the millennium development goal affects inequality between the poorest and the least poor? an analysis of demographic and health survey. *British Medical Journal*, 7526:1180–1182.

Mosley, W. H. and Chen, L. C. (1984). An analytical framework for the study of child survival in developing countries. *Population and Development Review*, 10(7):25–45.

Mpimbaza, A., Sears, D., Sserwanga, A., Kigozi, R., Rubahika, D., Nadler, A., Yeka, A., and Dorsey, G. (2015). Admission risk score to predict inpatient pediatric mortality at four public hospitals in uganda. *PLoS ONE*, 10(7):e0133950.

Rajaratnam, J. K., Marcus, J. R., Flaxman, A. D., Wang, H., Levin-Rector, A., Dwyer, L., Costa, M., Lopez, A. D., and Murray, C. J. L. (2010). Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards millennium development goal 4. *The Lancet*.

Ramos, A. P., Heymann, J., and Weiss, R. (2018). Improving program targeting to combat early-life mortality by identifying high risk births in india. *Population Health Metrics*, 16(17).

Ramos, A. P., Weiss, R. E., and Flores, M. (2020). Leave no child behind: Using data from 1.7 million children from 67 developing countries to measure inequality within and between groups of births and to identify left behind populations. *PlosOne*, (7):25–45.

Stuckler, D., Basu, S., and Mckee, M. (2010). Drivers of inequality in Millennium Development Goal progress: a statistical analysis. *PLoS Medicine*, 7(3):e1000241.

World Bank Development Indicators (2021). Birth rates, early mortality. Data retrieved from World Development Indicators, `http://data.worldbank.org/`.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67:301–320.

A Appendix

A.1 Cross validation results

Table 5: Recall10 results, cross-validation on training set

Survey	wealth	elastic-logit	rf	xgb	krls	ensemble	efficiency gain
Angola 2015	0.11	0.21	0.21	0.22	0.24	0.24	2.30
Benin 2011	0.12	0.28	0.26	0.28	0.29	0.29	2.30
Burkina Faso 2010	0.09	0.16	0.15	0.16	0.14	0.18	2.08
Cameroon 2011	0.12	0.15	0.22	0.14	0.15	0.24	2.04
Congo Democratic Republic 2013-14	0.09	0.14	0.19	0.16	0.17	0.20	2.27
Cote d'Ivoire 2011	0.11	0.20	0.22	0.19	0.21	0.21	1.94
Ghana 2014	0.06	0.18	0.13	0.17	0.16	0.18	3.11
Guinea 2012	0.17	0.16	0.15	0.14	0.16	0.17	1.00
Kenya 2014	0.08	0.15	0.17	0.13	0.15	0.16	2.14
Lesotho 2014	0.07	0.09	0.16	0.08	0.15	0.14	2.00
Madagascar 2008	0.10	0.24	0.24	0.24	0.25	0.27	2.66
Malawi 2016	0.11	0.16	0.20	0.18	0.20	0.21	1.94
Mali 2012	0.12	0.20	0.22	0.21	0.21	0.24	2.11
Mozambique 2011	0.10	0.16	0.24	0.18	0.21	0.24	2.34
Niger 2012	0.10	0.15	0.20	0.16	0.15	0.19	2.02
Nigeria 2013	0.12	0.17	0.17	0.19	0.19	0.20	1.60
Rwanda 2014	0.15	0.16	0.21	0.13	0.14	0.19	1.25
Senegal 2017	0.08	0.15	0.21	0.21	0.17	0.22	2.67
Tanzania 2015	0.09	0.12	0.18	0.16	0.14	0.19	2.18
Uganda 2016	0.10	0.14	0.17	0.16	0.13	0.18	1.73
Zambia 2013	0.08	0.15	0.21	0.19	0.19	0.21	2.61
Zimbabwe 2015	0.12	0.14	0.20	0.17	0.15	0.19	1.57
Average	0.10	0.17	0.20	0.18	0.18	0.21	2.08

Table 6: Recall20 results, cross-validation on training set

Survey	wealth	elastic-logit	rf	xgb	krls	ensemble	efficiency gain
Angola 2015	0.24	0.39	0.37	0.40	0.38	0.42	1.76
Benin 2011	0.24	0.45	0.41	0.45	0.45	0.47	1.97
Burkina Faso 2010	0.23	0.31	0.29	0.31	0.30	0.30	1.30
Cameroon 2011	0.25	0.27	0.33	0.27	0.27	0.37	1.45
Congo Democratic Republic 2013-14	0.18	0.26	0.30	0.29	0.29	0.32	1.76
Cote d'Ivoire 2011	0.20	0.32	0.35	0.32	0.33	0.39	1.99
Ghana 2014	0.21	0.30	0.24	0.28	0.29	0.27	1.28
Guinea 2012	0.28	0.27	0.29	0.30	0.27	0.32	1.12
Kenya 2014	0.17	0.27	0.28	0.27	0.25	0.30	1.80
Lesotho 2014	0.14	0.22	0.32	0.18	0.25	0.27	1.94
Madagascar 2008	0.22	0.38	0.36	0.39	0.39	0.42	1.90
Malawi 2016	0.22	0.30	0.31	0.32	0.32	0.35	1.63
Mali 2012	0.22	0.36	0.35	0.36	0.38	0.37	1.72
Mozambique 2011	0.24	0.29	0.37	0.33	0.32	0.37	1.53
Niger 2012	0.20	0.31	0.31	0.29	0.28	0.33	1.71
Nigeria 2013	0.25	0.34	0.29	0.34	0.33	0.34	1.35
Rwanda 2014	0.26	0.36	0.30	0.27	0.31	0.32	1.23
Senegal 2017	0.20	0.30	0.31	0.33	0.31	0.34	1.68
Tanzania 2015	0.16	0.23	0.31	0.29	0.25	0.26	1.63
Uganda 2016	0.22	0.24	0.28	0.29	0.24	0.31	1.37
Zambia 2013	0.16	0.31	0.30	0.32	0.33	0.35	2.17
Zimbabwe 2015	0.23	0.31	0.34	0.32	0.29	0.33	1.47
Average	0.22	0.31	0.32	0.31	0.31	0.34	1.62

A.2 Variable construction

Table 7: Variable information

Variable	Type	Range of Values	Description
kidbordlog	Numeric	[0.6931,2.9444]	Natural logarithm of the child's birth order
pregtermin	Categorical	{NoPregTerm, PregTerm}	Ever had pregnancy terminate
agefrstmar	Numeric	[1,49]	Women's age at start of first marriage or union
hheadagehh	Numeric	[8,95]	Age of household head
urban	Categorical	{Urban, Rural}	Urban or rural living status
kidsex	Categorical	{girl, boy}	Sex of child
hheadsex	Categorical	{female, male}	Sex of household head
edyrtotal	Numeric	[1,25]	Total years educated
maternal_age_month	Numeric	[135,584]	Age of mother calculated in months
drinkwtr_new	Categorical	{bad, good}	Main source of drinking water
floor_new	Categorical	{unsafe, safe}	Main material of floor
cookfuel_new	Categorical	{unclean, clean}	Type of fuel used for cooking
toilettype_new	Categorical	{flush, pit, unimproved}	Type of toilet
religion_new	Categorical	{Buddhist, Christian, Hindu, Other}	Religion of mother
wealths	Numeric	[-16.3862,32.4577]	Wealth index categorical score
kidbirthmo	Categorical	{1,2,âŠ,12}	Month of the child's birth
kidbirthmo_sin	Numeric	[-1,1]	Sine transformation of child's birth month
kidbirthmo_cos	Numeric	[-1,1]	Cosine transformation of child's birth month
bednetnum	Numeric	{0,1,2,3,4,5,6,7+}	Number of mosquito nets owned by the household
bednetnum_log	Numeric	[0.69,2.2]	Natural logarithm of bednetnum+1
prev_death	Categorical	{no, yes}	Mother experienced previous death of child ¡12mo
prev_death_full	Categorical	{no, yes}	Mother experienced previous death of child ¡5yr
district	Categorical	Depends on the country	District where the household lived at the time of survey
Province/State	Categorical	Depends on the country	State or Province where the household lived at the time of survey
wealthp2	Categorical	{1,2,3,4,5}	Wealth quintile based on wealths (by country)

B Online Supplement

B.1 Country-by-Country Results

Efficiency gain is measured by the ratio of Catch 10 Results for each algorithm compared to the Catch 10 Results for using wealth alone.

Table 8: Detailed Results by Country

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
Angola							
Elastic Net	0.212	0.394	0.664	0.012	1.276	0.933	2.001
Ensemble	0.244	0.424	0.683	0.014	1.399	0.934	2.300
KRLS	0.236	0.384	0.669	0.016	1.379	0.934	2.227
Random Forest	0.214	0.368	0.641	0.013	1.904	0.933	2.024
Wealth	0.106	0.241	0.560	0.207	15.020	0.928	1.000
XGB	0.225	0.397	0.666	0.016	1.379	0.934	2.124
Benin							
Elastic Net	0.284	0.454	0.678	0.026	1.568	0.936	2.276
Ensemble	0.287	0.475	0.685	0.023	1.590	0.935	2.299
KRLS	0.292	0.449	0.674	0.021	1.475	0.936	2.338
Random Forest	0.260	0.411	0.662	0.022	2.046	0.935	2.083
Wealth	0.125	0.242	0.537	0.118	2.244	0.928	1.000
XGB	0.278	0.446	0.666	0.024	1.518	0.935	2.233
Burkina Faso							
Elastic Net	0.161	0.307	0.608	0.008	1.120	0.920	1.866
Ensemble	0.179	0.304	0.626	0.010	1.175	0.921	2.076

Table 8: Detailed Results by Country *(continued)*

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
KRLS	0.145	0.297	0.611	0.010	1.152	0.919	1.673
Random Forest	0.155	0.287	0.581	0.009	1.349	0.920	1.789
Wealth	0.086	0.234	0.553	0.147	1.958	0.914	1.000
XGB	0.159	0.306	0.620	0.013	1.188	0.920	1.845
Cameroon							
Elastic Net	0.153	0.272	0.584	0.006	1.085	0.921	1.304
Ensemble	0.239	0.365	0.620	0.021	1.313	0.926	2.040
KRLS	0.146	0.266	0.584	0.005	1.074	0.920	1.247
Random Forest	0.217	0.327	0.598	0.062	1.827	0.925	1.850
Wealth	0.117	0.252	0.547	0.137	2.072	0.918	1.000
XGB	0.142	0.274	0.595	0.013	1.199	0.920	1.210
Congo Democratic Republic							
Elastic Net	0.145	0.263	0.557	0.004	1.064	0.922	1.625
Ensemble	0.202	0.325	0.604	0.008	1.162	0.926	2.266
KRLS	0.174	0.293	0.588	0.008	1.126	0.924	1.945
Random Forest	0.190	0.296	0.581	0.015	1.618	0.925	2.124
Wealth	0.089	0.185	0.520	0.064	1.503	0.919	1.000
XGB	0.162	0.294	0.590	0.007	1.116	0.923	1.821
Cote d'Ivoire							
Elastic Net	0.205	0.321	0.629	0.016	1.205	0.920	1.918
Ensemble	0.208	0.395	0.654	0.018	1.241	0.919	1.943
KRLS	0.208	0.330	0.638	0.016	1.205	0.920	1.947

Table 8: Detailed Results by Country *(continued)*

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
Random Forest	0.217	0.353	0.622	0.024	1.432	0.921	2.027
Wealth	0.107	0.199	0.529	0.076	2.132	0.911	1.000
XGB	0.193	0.324	0.619	0.015	1.185	0.919	1.806
Ghana							
Elastic Net	0.185	0.304	0.519	0.002	1.048	0.932	3.127
Ensemble	0.185	0.271	0.539	0.007	1.161	0.931	3.127
KRLS	0.165	0.290	0.567	0.000	1.009	0.931	2.796
Random Forest	0.132	0.243	0.556	0.019	1.530	0.929	2.232
Wealth	0.059	0.210	0.568	-0.038	0.889	0.926	1.000
XGB	0.172	0.285	0.539	0.006	1.135	0.931	2.901
Guinea							
Elastic Net	0.157	0.274	0.576	0.005	1.070	0.918	0.926
Ensemble	0.170	0.318	0.606	0.009	1.134	0.918	1.001
KRLS	0.157	0.274	0.582	0.005	1.071	0.918	0.925
Random Forest	0.145	0.287	0.582	0.012	1.301	0.917	0.853
Wealth	0.170	0.283	0.574	0.242	2.981	0.919	1.000
XGB	0.139	0.296	0.599	0.013	1.171	0.916	0.819
Kenya							
Elastic Net	0.148	0.271	0.524	0.001	1.025	0.932	1.948
Ensemble	0.163	0.301	0.573	0.003	1.088	0.932	2.143
KRLS	0.150	0.250	0.536	0.000	1.005	0.932	1.975
Random Forest	0.168	0.282	0.564	0.009	1.408	0.933	2.199

Table 8: Detailed Results by Country *(continued)*

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
Wealth	0.076	0.167	0.542	-0.028	0.882	0.929	1.000
XGB	0.127	0.273	0.558	0.002	1.046	0.931	1.668
Lesotho							
Elastic Net	0.094	0.214	0.575	0.001	1.011	0.916	1.378
Ensemble	0.138	0.267	0.586	0.005	1.076	0.918	2.022
KRLS	0.145	0.249	0.548	0.001	1.017	0.920	2.133
Random Forest	0.162	0.317	0.582	0.013	1.344	0.921	2.378
Wealth	0.068	0.137	0.583	-0.026	0.859	0.914	1.000
XGB	0.077	0.181	0.522	0.003	1.047	0.915	1.133
Madagascar							
Elastic Net	0.242	0.378	0.629	0.013	1.270	0.933	2.339
Ensemble	0.274	0.424	0.656	0.024	1.473	0.934	2.652
KRLS	0.250	0.391	0.642	0.014	1.288	0.933	2.417
Random Forest	0.236	0.361	0.635	0.052	1.933	0.932	2.285
Wealth	0.103	0.223	0.549	0.126	1.532	0.925	1.000
XGB	0.245	0.391	0.653	0.016	1.338	0.933	2.364
Malawi							
Elastic Net	0.164	0.301	0.582	0.004	1.086	0.931	1.508
Ensemble	0.211	0.350	0.619	0.009	1.238	0.933	1.937
KRLS	0.202	0.321	0.594	0.006	1.153	0.933	1.857
Random Forest	0.204	0.310	0.600	0.017	1.910	0.933	1.878
Wealth	0.109	0.215	0.533	0.063	1.456	0.929	1.000

Table 8: Detailed Results by Country *(continued)*

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
XGB	0.184	0.319	0.594	0.007	1.175	0.932	1.692
Mali							
Elastic Net	0.203	0.355	0.646	0.015	1.262	0.927	1.750
Ensemble	0.245	0.371	0.666	0.025	1.410	0.929	2.114
KRLS	0.213	0.379	0.657	0.018	1.299	0.927	1.841
Random Forest	0.221	0.347	0.634	0.050	1.742	0.928	1.909
Wealth	0.116	0.216	0.555	0.181	3.474	0.921	1.000
XGB	0.208	0.358	0.654	0.017	1.291	0.927	1.795
Mozambique							
Elastic Net	0.156	0.286	0.567	0.005	1.077	0.921	1.523
Ensemble	0.240	0.370	0.631	0.020	1.296	0.926	2.341
KRLS	0.207	0.323	0.591	0.010	1.148	0.924	2.023
Random Forest	0.244	0.365	0.625	0.051	1.676	0.927	2.386
Wealth	0.102	0.242	0.536	0.089	2.105	0.917	1.000
XGB	0.179	0.326	0.598	0.015	1.228	0.922	1.750
Niger							
Elastic Net	0.149	0.307	0.612	0.008	1.127	0.923	1.554
Ensemble	0.194	0.335	0.635	0.018	1.287	0.925	2.021
KRLS	0.149	0.282	0.625	0.009	1.154	0.923	1.554
Random Forest	0.205	0.307	0.610	0.046	1.671	0.927	2.131
Wealth	0.096	0.196	0.532	0.158	2.954	0.919	1.000
XGB	0.160	0.286	0.607	0.009	1.147	0.924	1.665

Table 8: Detailed Results by Country *(continued)*

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
Nigeria							
Elastic Net	0.173	0.338	0.621	0.012	1.171	0.920	1.401
Ensemble	0.199	0.339	0.634	0.017	1.227	0.922	1.604
KRLS	0.186	0.332	0.623	0.012	1.164	0.921	1.502
Random Forest	0.175	0.294	0.603	0.031	1.375	0.920	1.413
Wealth	0.124	0.251	0.568	0.223	1.934	0.916	1.000
XGB	0.189	0.342	0.622	0.012	1.175	0.921	1.526
Rwanda							
Elastic Net	0.164	0.357	0.623	0.004	1.114	0.936	1.103
Ensemble	0.186	0.320	0.618	0.005	1.172	0.936	1.251
KRLS	0.142	0.312	0.609	0.003	1.099	0.935	0.952
Random Forest	0.207	0.296	0.578	0.010	1.517	0.938	1.387
Wealth	0.149	0.259	0.561	0.196	5.841	0.936	1.000
XGB	0.126	0.266	0.556	0.003	1.090	0.935	0.845
Senegal							
Elastic Net	0.149	0.299	0.580	0.004	1.101	0.930	1.773
Ensemble	0.224	0.339	0.609	0.033	1.810	0.933	2.660
KRLS	0.165	0.315	0.581	0.004	1.100	0.931	1.957
Random Forest	0.209	0.311	0.587	0.059	2.194	0.933	2.475
Wealth	0.084	0.203	0.535	0.106	1.687	0.927	1.000
XGB	0.209	0.333	0.595	0.064	3.297	0.933	2.476
Tanzania							

Table 8: Detailed Results by Country *(continued)*

Algorithm	Recall 10	Recall 20	ROC	MRD	MRR	F1	Efficiency Gain
Elastic Net	0.121	0.226	0.561	0.002	1.054	0.930	1.397
Ensemble	0.188	0.262	0.582	0.005	1.145	0.932	2.169
KRLS	0.144	0.254	0.549	0.004	1.098	0.931	1.667
Random Forest	0.180	0.309	0.584	0.011	1.459	0.932	2.080
Wealth	0.086	0.161	0.545	-0.128	0.251	0.928	1.000
XGB	0.164	0.285	0.560	0.005	1.110	0.932	1.895
Uganda							
Elastic Net	0.145	0.240	0.536	0.001	1.019	0.930	1.402
Ensemble	0.178	0.307	0.579	0.002	1.059	0.932	1.725
KRLS	0.134	0.237	0.535	0.001	1.027	0.930	1.299
Random Forest	0.167	0.278	0.554	0.001	1.620	0.931	1.624
Wealth	0.103	0.224	0.537	-0.021	0.919	0.929	1.000
XGB	0.157	0.294	0.576	0.004	1.104	0.931	1.526
Zambia							
Elastic Net	0.150	0.310	0.579	0.005	1.099	0.928	1.875
Ensemble	0.209	0.354	0.617	0.008	1.190	0.930	2.619
KRLS	0.186	0.333	0.588	0.005	1.098	0.930	2.327
Random Forest	0.209	0.300	0.598	0.016	1.588	0.931	2.618
Wealth	0.080	0.163	0.529	-0.009	0.990	0.924	1.000
XGB	0.194	0.320	0.595	0.008	1.160	0.930	2.422
Zimbabwe							
Elastic Net	0.143	0.308	0.576	0.005	1.095	0.926	1.176

Table 8: Detailed Results by Country *(continued)*

Algorithm	**Recall 10**	**Recall 20**	**ROC**	**MRD**	**MRR**	**F1**	**Efficiency Gain**
Ensemble	0.190	0.333	0.642	0.010	1.216	0.929	1.569
KRLS	0.154	0.292	0.600	0.005	1.108	0.927	1.265
Random Forest	0.201	0.338	0.629	0.019	1.726	0.930	1.653
Wealth	0.121	0.228	0.570	2.029	-170.854	0.925	1.000
XGB	0.175	0.318	0.588	0.009	1.184	0.928	1.441

B.2 Angola

This sample was taken in the year 2015. There are 10829 observations in the dataset, and 8664 observations were used for the training set. Also, 195 observations were removed due to missing values. Finally, there were 472 deaths for the full dataset, and 378 deaths in the training set.

B.2.1 Variables Used

Table 9: Variables included in the model for Angola

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_ao2015	categorical
wealthp	numeric
wealthp2	numeric

B.2.2 Mortality Breakdown by Wealth Quintile

B.2.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 10: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.55	0.0040056

Table 11: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 12: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	0.625

Table 13: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	80	4.85e-05

B.2.4 Table of Results

Table 14: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.244	0.424	0.683	0.099	0.014	1.399	0.041	0.934	2.300
KRLS	0.236	0.384	0.669	0.085	0.016	1.379	0.041	0.934	2.227
Elastic Net	0.212	0.394	0.664	0.083	0.012	1.276	0.041	0.933	2.001
Random Forest	0.214	0.368	0.641	0.102	0.013	1.904	0.042	0.933	2.024
Wealth	0.106	0.241	0.560	0.055	0.207	15.020	0.921	0.928	1.000
XGB	0.225	0.397	0.666	0.087	0.016	1.379	0.041	0.934	2.124

Table 15: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.153	0.166	0.173	0.144	0.153
2	0.163	0.177	0.186	0.110	0.195
3	0.140	0.143	0.152	0.097	0.131
4	0.134	0.149	0.140	0.127	0.158
5	0.115	0.126	0.115	0.150	0.134
6	0.108	0.087	0.090	0.114	0.076
7	0.067	0.060	0.053	0.064	0.053
8	0.050	0.040	0.040	0.060	0.036
9	0.042	0.033	0.030	0.066	0.031
10	0.027	0.026	0.020	0.067	0.030

B.2.5 Performance Plots

Below are various plots showing the performance for each model considered in Angola.

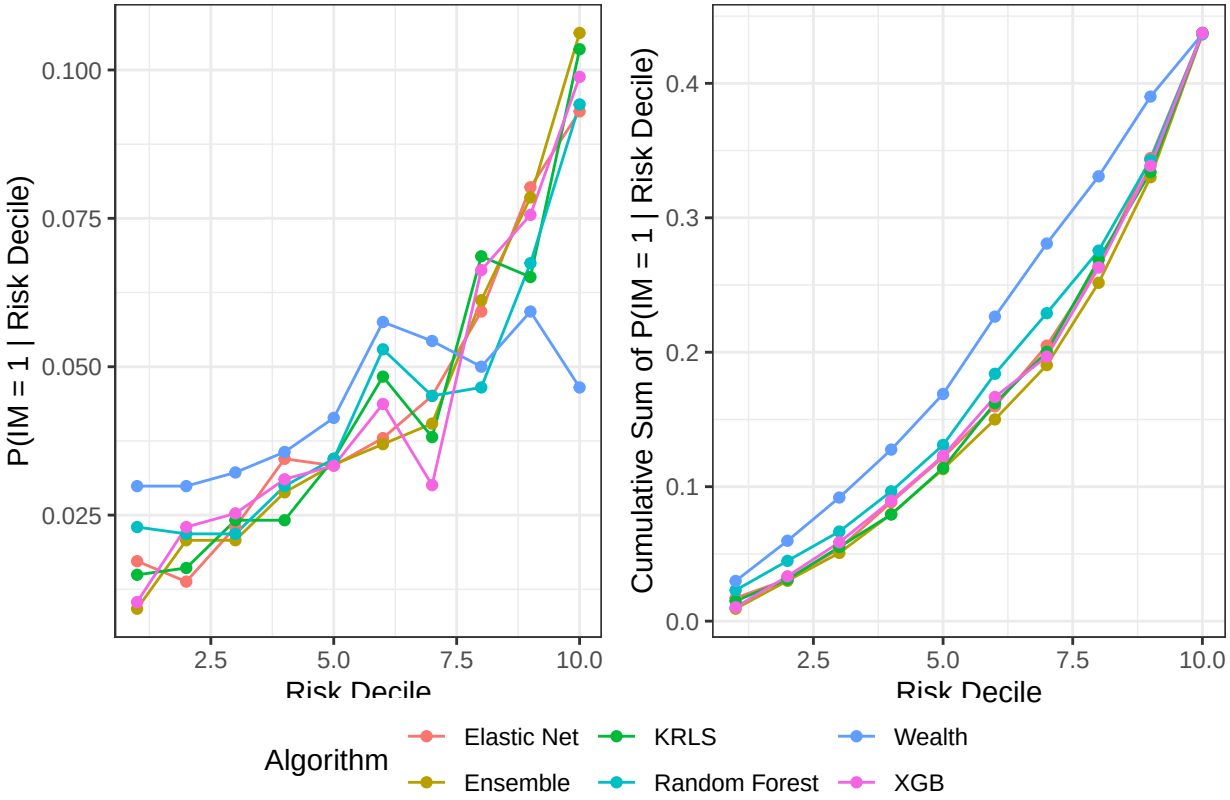

Figure 1: Probability of a mortality given that the observation is in a particular risk decile

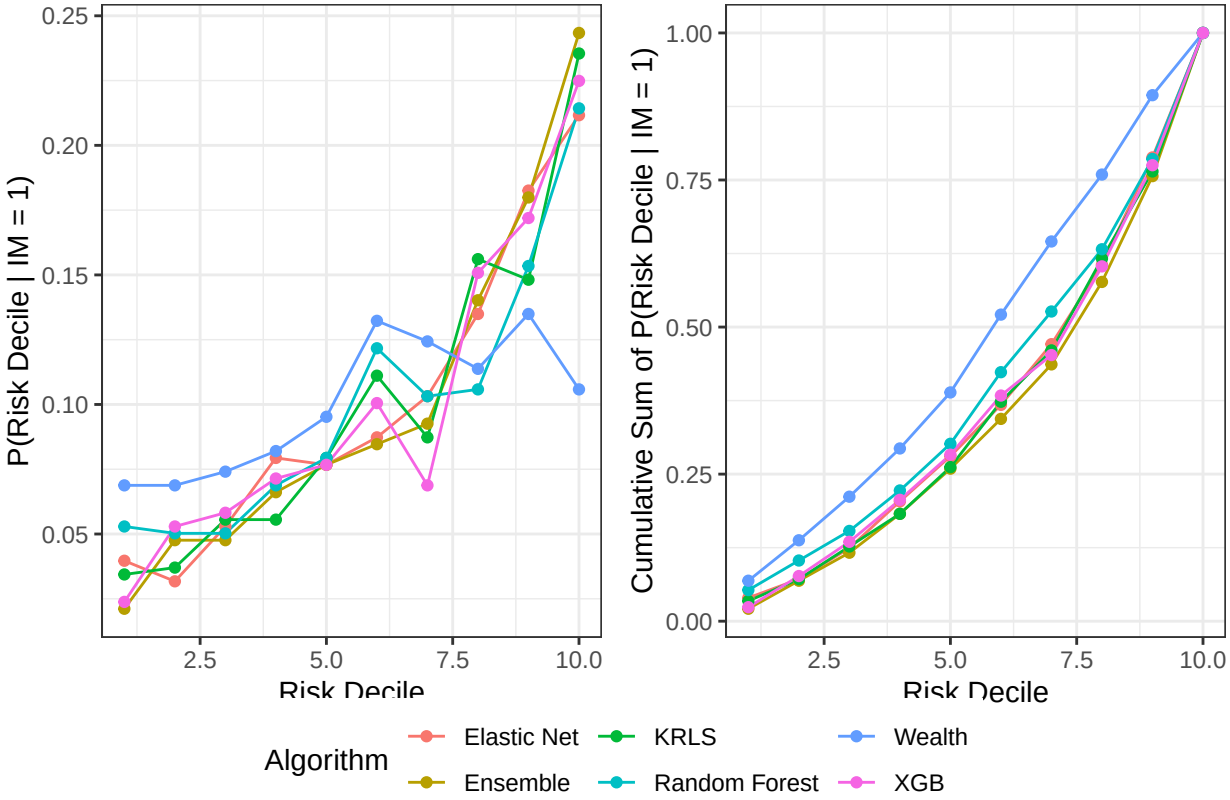

Figure 2: Probability of membership in a particular risk decile given a mortality

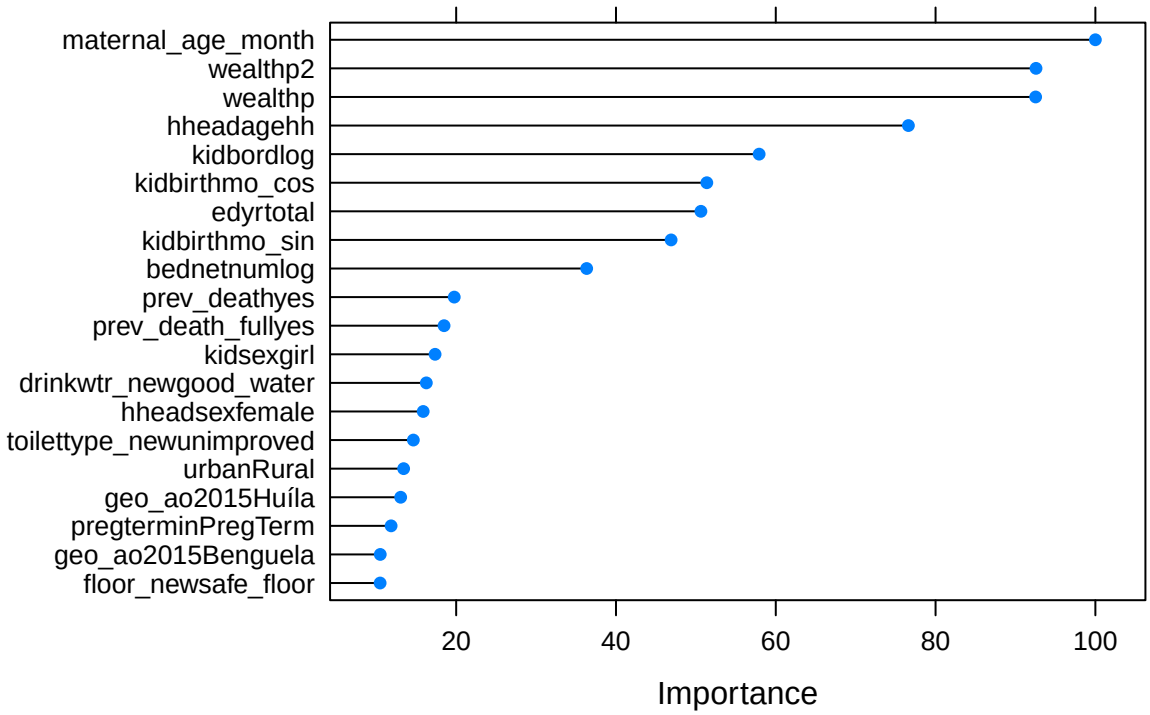

Figure 3: Variable Importance Plot generated by the Random Forest model

B.2.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.2.7 Elastic Net Coefficients

Table 16: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.5213	-2.5213
pregterminPregTerm	-	-
hheadagehh	6e-04	0.0073
urbanRural	-	-
kidsexgirl	-0.2118	-0.1059
hheadsexfemale	-	-
edyrtotal	-	-
drinkwtr_newgood_water	0.0334	0.016
floor_newsafe_floor	-0.0109	-0.0053
maternal_age_month	-0.002	-0.1704
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-	-
prev_deathyes	0.6642	0.2278
prev_death_fullyes	0.3219	0.1231
kidbordlog	-	-
kidbirthmo_sin	-	-
kidbirthmo_cos	-	-
geo_ao2015Zaire	-0.0145	-0.0032
geo_ao2015Uíge	-	-
geo_ao2015Luanda	-	-
geo_ao2015Cuanza Norte	-	-
geo_ao2015Cuanza Sul	0.0993	0.0228
geo_ao2015Malanje	-	-
geo_ao2015Lunda Norte	0.223	0.0504
geo_ao2015Benguela	0.3073	0.0747
geo_ao2015Huambo	0.1606	0.0405
geo_ao2015Bié	-	-
geo_ao2015Moxico	-0.5247	-0.0982
geo_ao2015Cuando Cubango	-	-
geo_ao2015Namibe	-	-
geo_ao2015Huíla	0.389	0.0965
geo_ao2015Cunene	-	-
geo_ao2015Lunda Sul	-	-

Table 16: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation. *(continued)*

	Coefficients	**Standardized Coefficients**
geo_ao2015Bengo	-0.2043	-0.041
wealthp	-	-
wealthp2	-0.4321	-0.129

B.3 Benin

This sample was taken in the year 2011. There are 10199 observations in the dataset, and 8160 observations were used for the training set. Also, 291 observations were removed due to missing values. Finally, there were 471 deaths for the full dataset, and 377 deaths in the training set.

B.3.1 Variables Used

Table 17: Variables included in the model for Benin

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_bj2011	categorical
wealthp	numeric
wealthp2	numeric

B.3.2 Mortality Breakdown by Wealth Quintile

B.3.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 18: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.0063933

Table 19: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 20: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	0.75

Table 21: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	72	0.0001239

B.3.4 Table of Results

Table 22: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.287	0.475	0.685	0.109	0.023	1.590	0.043	0.935	2.299
KRLS	0.292	0.449	0.674	0.098	0.021	1.475	0.043	0.936	2.338
Elastic Net	0.284	0.454	0.678	0.106	0.026	1.568	0.043	0.936	2.276
Random Forest	0.260	0.411	0.662	0.123	0.022	2.046	0.044	0.935	2.083
Wealth	0.125	0.242	0.537	0.061	0.118	2.244	0.894	0.928	1.000
XGB	0.278	0.446	0.666	0.097	0.024	1.518	0.043	0.935	2.233

Table 23: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.141	0.154	0.147	0.165	0.146
2	0.140	0.142	0.138	0.124	0.152
3	0.142	0.147	0.143	0.110	0.137
4	0.147	0.141	0.147	0.102	0.140
5	0.077	0.078	0.081	0.088	0.077
6	0.094	0.099	0.099	0.094	0.098
7	0.091	0.085	0.086	0.067	0.072
8	0.068	0.059	0.067	0.068	0.069
9	0.062	0.062	0.059	0.098	0.068
10	0.038	0.040	0.031	0.083	0.041

B.3.5 Performance Plots

Below are various plots showing the performance for each model considered in Benin.

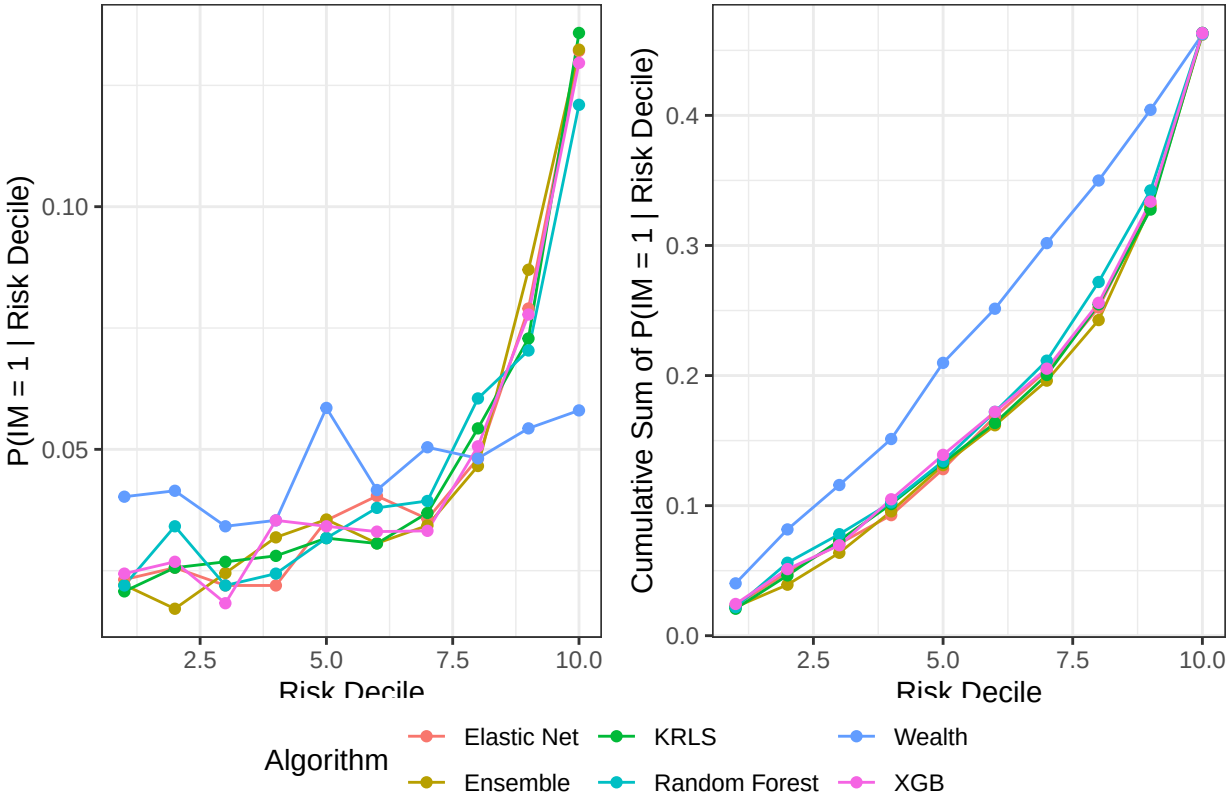

Figure 4: Probability of a mortality given that the observation is in a particular risk decile

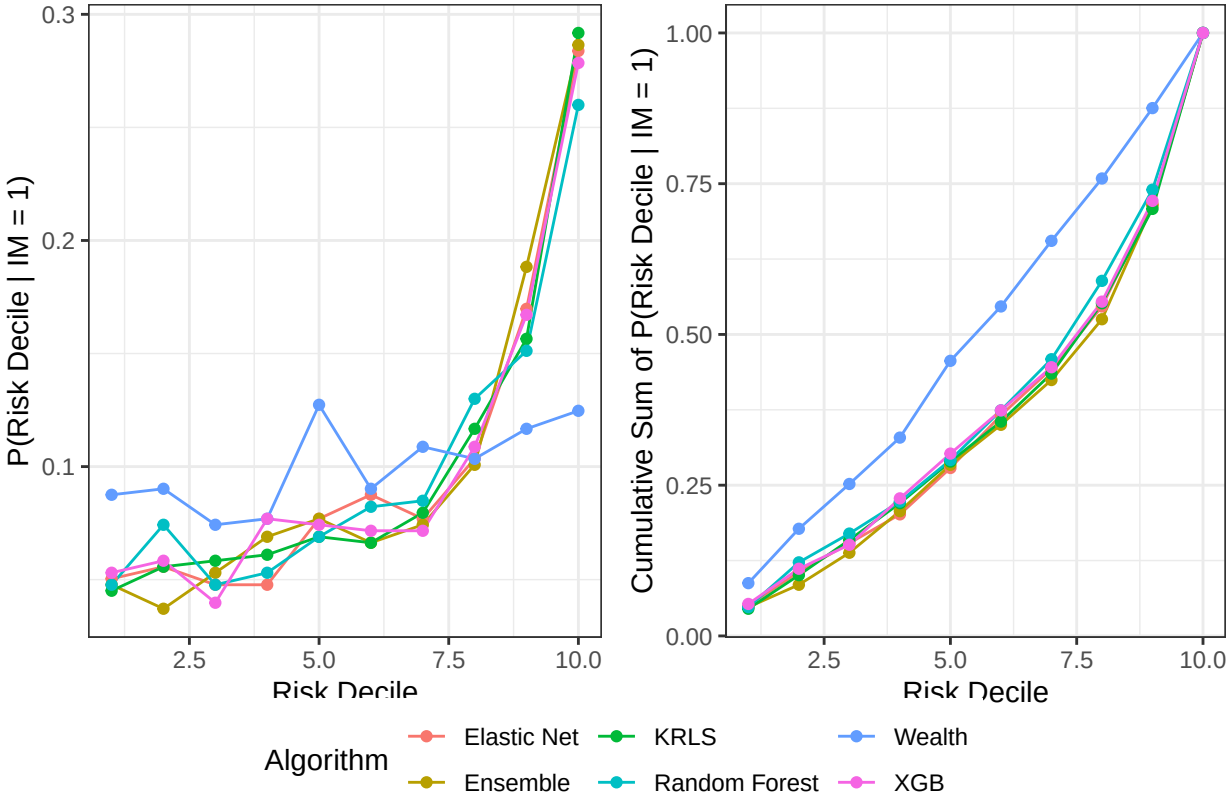

Figure 5: Probability of membership in a particular risk decile given a mortality

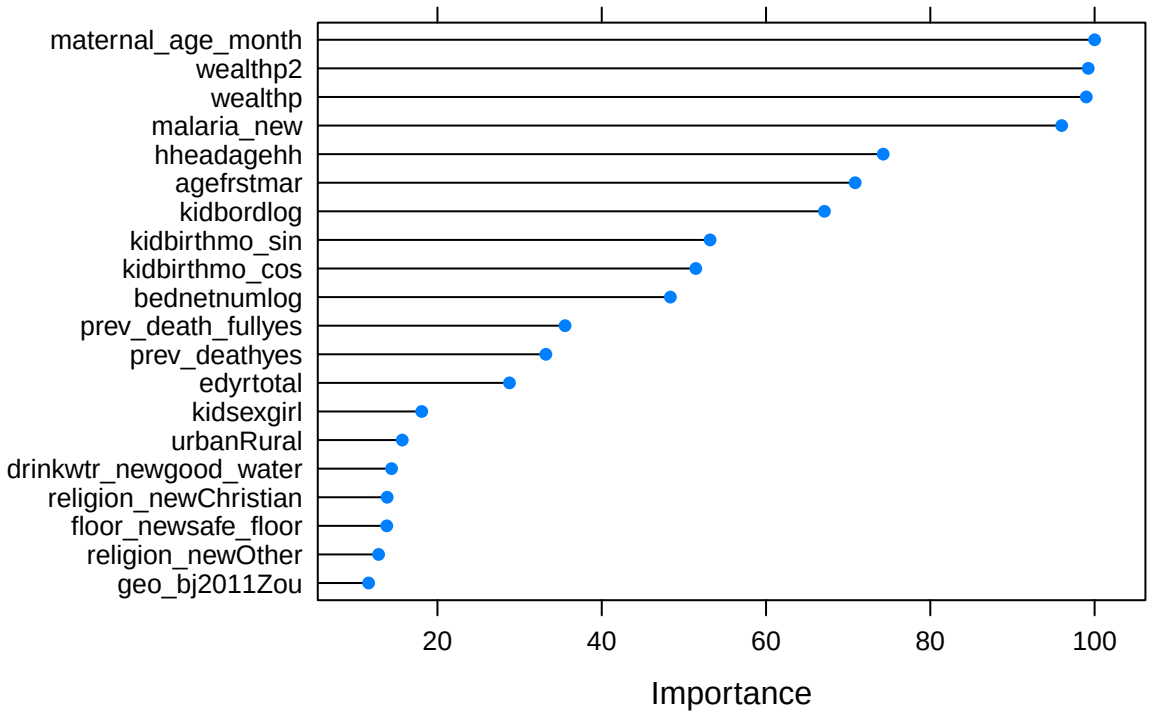

Figure 6: Variable Importance Plot generated by the Random Forest model

B.3.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.3.7 Elastic Net Coefficients

Table 24: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.2806	-2.2806
pregterminPregTerm	-0.2594	-0.0625
agefrstmar	-0.0112	-0.0498
hheadagehh	-8e-04	-0.009
urbanRural	-0.123	-0.0591
kidsexgirl	-0.1702	-0.0851
hheadsexfemale	-0.1426	-0.0448
edyrtotal	-0.0077	-0.0239
drinkwtr_newgood_water	0.0945	0.0412
floor_newsafe_floor	-0.1028	-0.0509
maternal_age_month	-0.0018	-0.1317
cookfuel_newclean_fuel	-	-
toilettype_newpit	-0.1219	-0.056
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-0.0057	-0.0025
religion_newOther	0.0498	0.0207
bednetnumlog	-0.0252	-0.0128
prev_deathyes	0.5222	0.1642
prev_death_fullyes	0.8644	0.3144
malaria_new	9e-04	1e-04
kidbordlog	0.1375	0.0896
kidbirthmo_sin	0.0954	0.0682
kidbirthmo_cos	-0.0308	-0.0216
geo_bj2011Atacora	-0.3459	-0.1056
geo_bj2011Atlantique	0.1161	0.0376
geo_bj2011Borgou	-0.2252	-0.0622
geo_bj2011Collines	-0.0098	-0.0025
geo_bj2011Couffo	-	-
geo_bj2011Donga	0.0855	0.0195
geo_bj2011Littoral	0.4842	0.1308
geo_bj2011Mono	-0.3927	-0.0926
geo_bj2011Ouémé	-0.3573	-0.1126
geo_bj2011Plateau	-0.4086	-0.1009
geo_bj2011Zou	0.0573	0.017
wealthp	-0.1813	-0.0524
wealthp2	-0.0293	-0.0088

B.4 Burkina Faso

This sample was taken in the year 2010. There are 10905 observations in the dataset, and 8725 observations were used for the training set. Also, 823 observations were removed due to missing values. Finally, there were 752 deaths for the full dataset, and 602 deaths in the training set.

B.4.1 Variables Used

Table 25: Variables included in the model for Burkina Faso

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_bf2010	categorical
wealthp	numeric
wealthp2	numeric

B.4.2 Mortality Breakdown by Wealth Quintile

B.4.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 26: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.0214685

Table 27: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 28: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	0.875

Table 29: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	74	0.0001417

B.4.4 Table of Results

Table 30: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.179	0.304	0.626	0.115	0.010	1.175	0.064	0.921	2.076
KRLS	0.145	0.297	0.611	0.111	0.010	1.152	0.064	0.919	1.673
Elastic Net	0.161	0.307	0.608	0.111	0.008	1.120	0.064	0.920	1.866
Random Forest	0.155	0.287	0.581	0.105	0.009	1.349	0.065	0.920	1.789
Wealth	0.086	0.234	0.553	0.077	0.147	1.958	0.720	0.914	1.000
XGB	0.159	0.306	0.620	0.112	0.013	1.188	0.064	0.920	1.845

Table 31: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.229	0.171	0.231	0.116	0.101
2	0.216	0.225	0.212	0.126	0.242
3	0.169	0.168	0.177	0.090	0.183
4	0.145	0.149	0.140	0.097	0.180
5	0.080	0.089	0.090	0.097	0.107
6	0.054	0.071	0.060	0.102	0.078
7	0.050	0.049	0.049	0.093	0.047
8	0.041	0.037	0.030	0.095	0.032
9	0.010	0.033	0.009	0.099	0.025
10	0.005	0.009	0.001	0.085	0.003

B.4.5 Performance Plots

Below are various plots showing the performance for each model considered in Burkina Faso.

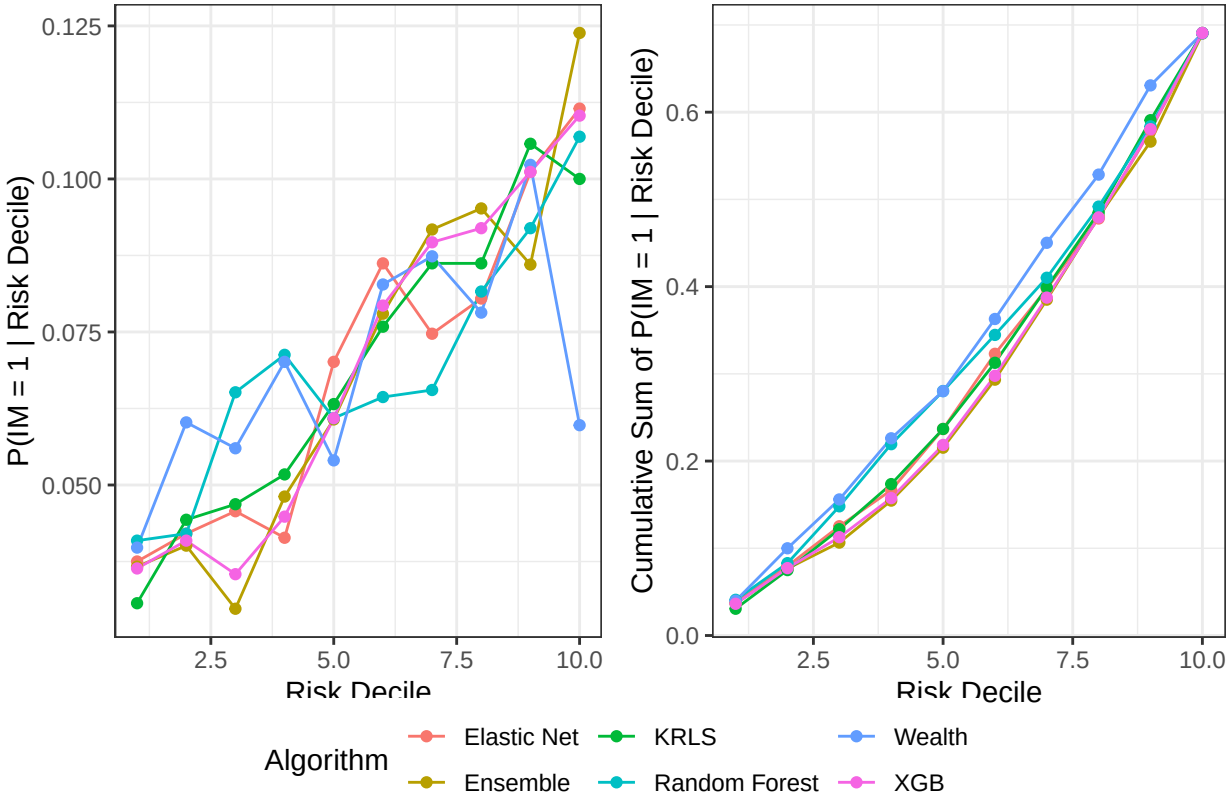

Figure 7: Probability of a mortality given that the observation is in a particular risk decile

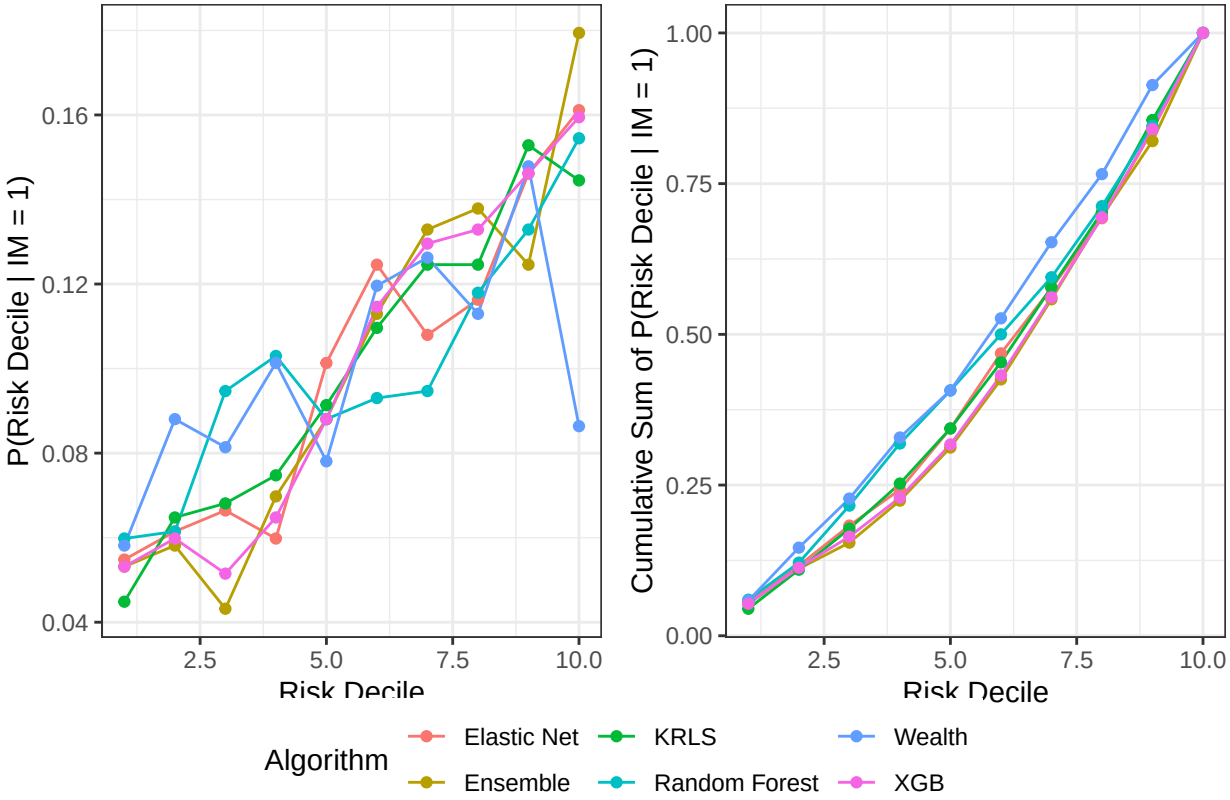

Figure 8: Probability of membership in a particular risk decile given a mortality

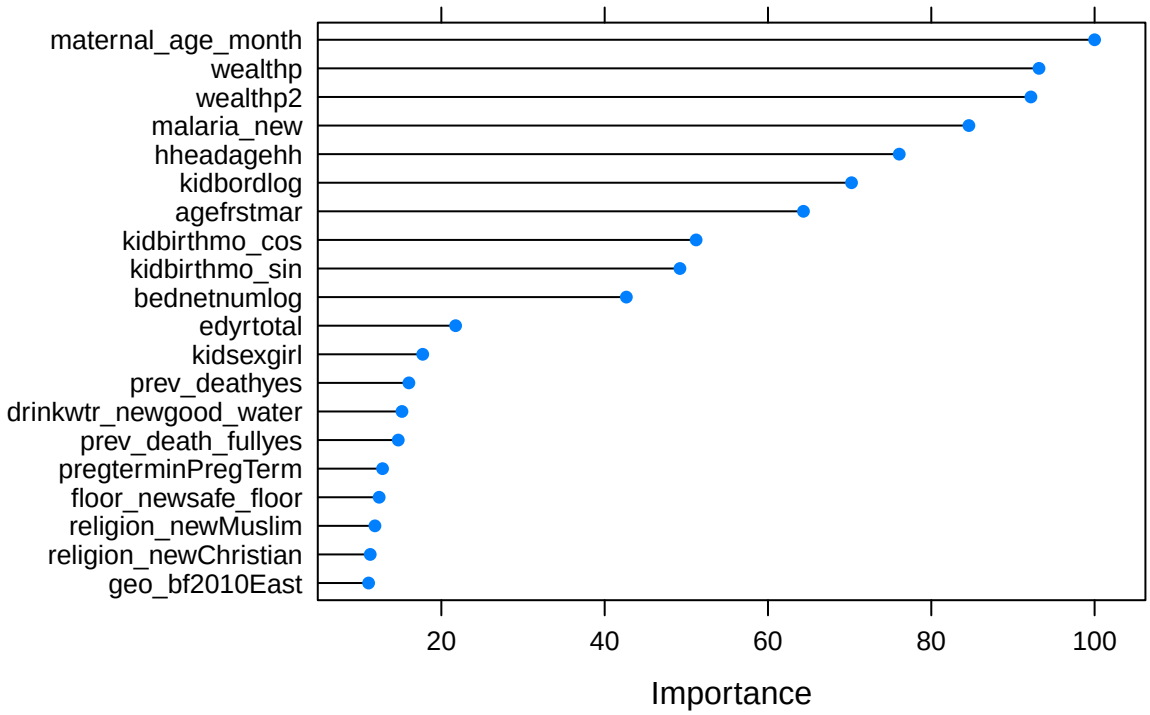

Figure 9: Variable Importance Plot generated by the Random Forest model

B.4.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.4.7 Elastic Net Coefficients

Table 32: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.4629	-2.4629
pregterminPregTerm	0.0037	0.0012
agefrstmar	-	-
hheadagehh	-	-
urbanRural	0.0872	0.036
kidsexgirl	-	-
hheadsexfemale	-	-
edyrtotal	-0.0146	-0.0363
drinkwtr_newgood_water	-0.147	-0.0632
floor_newsafe_floor	-	-
maternal_age_month	-7e-04	-0.0601
cookfuel_newclean_fuel	-0.1543	-0.0203
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	0.2417	0.0734
bednetnumlog	-	-
prev_deathyes	0.3052	0.1283
prev_death_fullyes	0.1653	0.0781
malaria_new	-	-
kidbordlog	-	-
kidbirthmo_sin	0.0234	0.0168
kidbirthmo_cos	-4e-04	-3e-04
geo_bf2010Cascades	-	-
geo_bf2010Center	-	-
geo_bf2010Center-East	-0.1013	-0.0276
geo_bf2010Center-North	-0.1151	-0.0304
geo_bf2010Center-West	-	-
geo_bf2010Center-South	-	-
geo_bf2010East	0.1602	0.0477
geo_bf2010Hauts Basins	-0.0208	-0.0058
geo_bf2010North	-	-
geo_bf2010Central Plateau	-0.057	-0.0137
geo_bf2010Sahel	0.1855	0.0531
geo_bf2010South-West	0.2369	0.0569
wealthp	-	-
wealthp2	-0.1237	-0.0369

B.5 Cameroon

This sample was taken in the year 2011. There are 8450 observations in the dataset, and 6761 observations were used for the training set. Also, 591 observations were removed due to missing values. Finally, there were 564 deaths for the full dataset, and 452 deaths in the training set.

B.5.1 Variables Used

Table 33: Variables included in the model for Cameroon

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_cm2011	categorical
wealthp	numeric
wealthp2	numeric

B.5.2 Mortality Breakdown by Wealth Quintile

B.5.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 34: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.0163135

Table 35: Optimal Parameters for the Random Forest algorithm

mtry
10

Table 36: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.4	0	0.8	1	0.75

Table 37: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	72	0.0004365

B.5.4 Table of Results

Table 38: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.239	0.365	0.620	0.182	0.021	1.313	0.061	0.926	2.040
KRLS	0.146	0.266	0.584	0.092	0.005	1.074	0.062	0.920	1.247
Elastic Net	0.153	0.272	0.584	0.095	0.006	1.085	0.062	0.921	1.304
Random Forest	0.217	0.327	0.598	0.179	0.062	1.827	0.061	0.925	1.850
Wealth	0.117	0.252	0.547	0.084	0.137	2.072	0.910	0.918	1.000
XGB	0.142	0.274	0.595	0.114	0.013	1.199	0.062	0.920	1.210

Table 39: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.272	0.218	0.246	0.188	0.204
2	0.212	0.186	0.223	0.145	0.237
3	0.128	0.109	0.131	0.104	0.116
4	0.106	0.078	0.116	0.064	0.091
5	0.085	0.107	0.094	0.094	0.097
6	0.072	0.085	0.073	0.091	0.082
7	0.050	0.074	0.055	0.075	0.047
8	0.022	0.039	0.024	0.046	0.033
9	0.036	0.052	0.021	0.079	0.030
10	0.015	0.058	0.013	0.110	0.060

B.5.5 Performance Plots

Below are various plots showing the performance for each model considered in Cameroon.

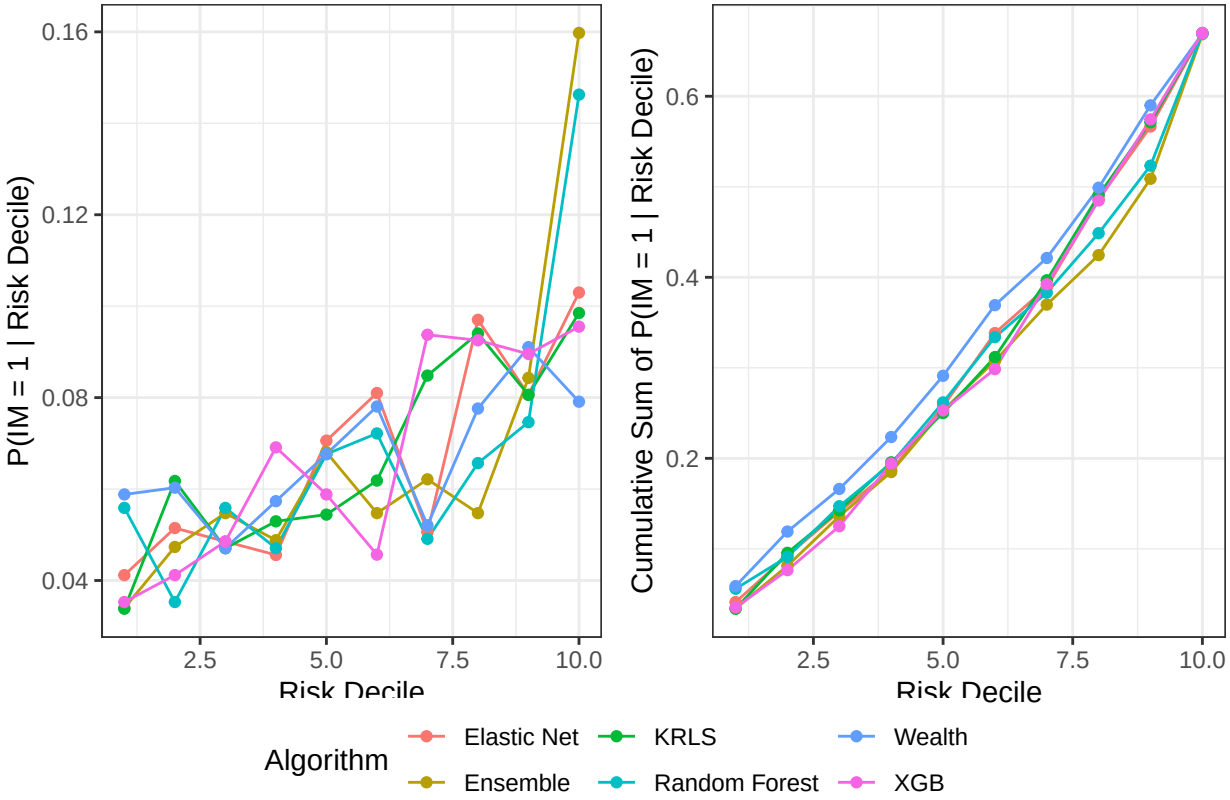

Figure 10: Probability of a mortality given that the observation is in a particular risk decile

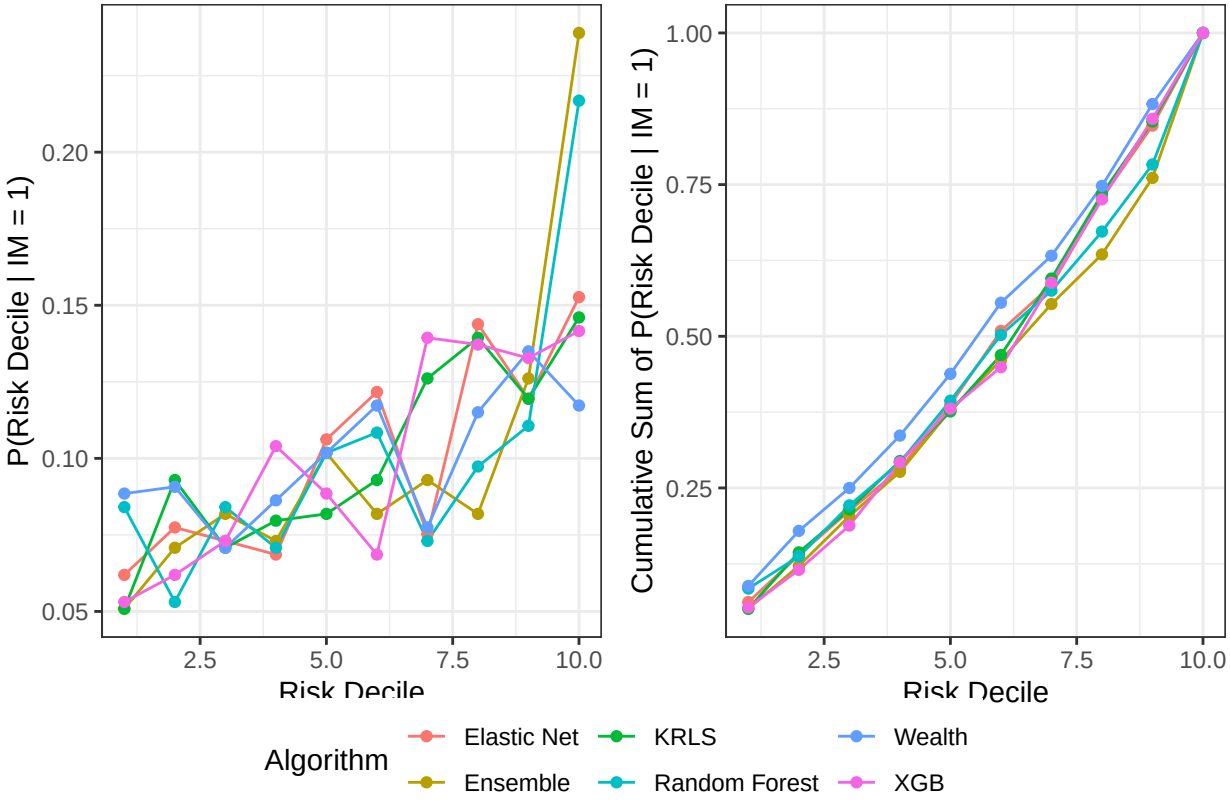

Figure 11: Probability of membership in a particular risk decile given a mortality

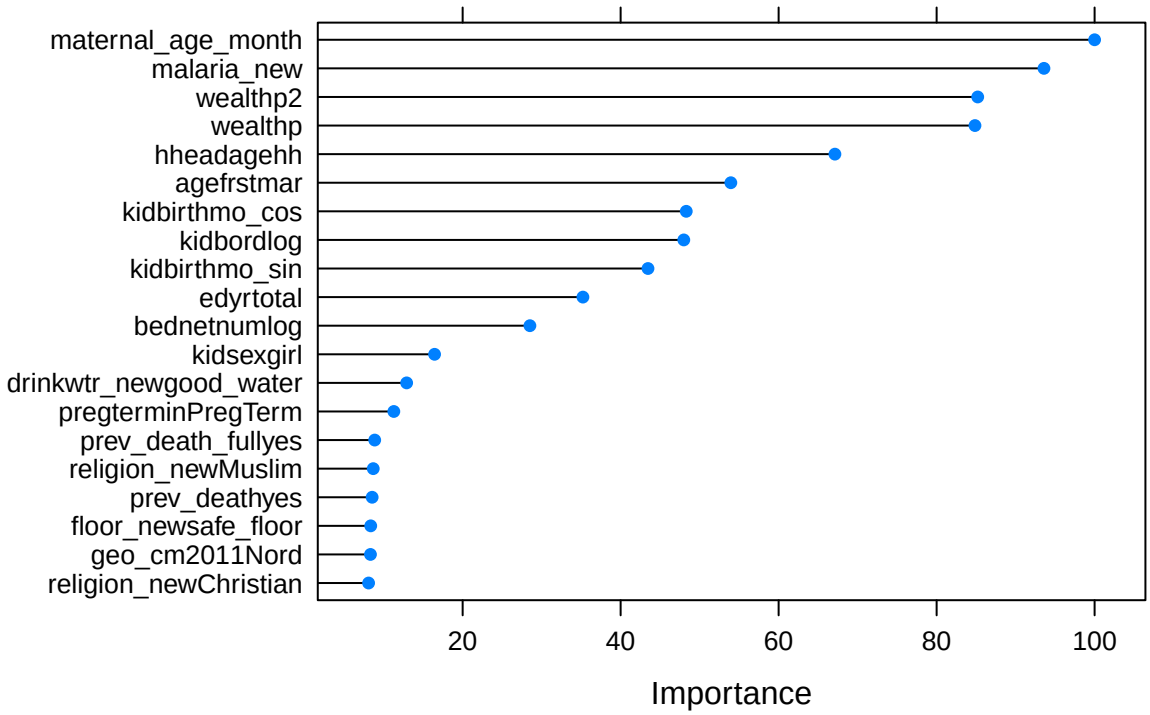

Figure 12: Variable Importance Plot generated by the Random Forest model

B.5.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.5.7 Elastic Net Coefficients

Table 40: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.3741	-2.3741
pregterminPregTerm	0.1506	0.0657
agefrstmar	-0.0041	-0.0162
hheadagehh	0.0036	0.0483
urbanRural	-	-
kidsexgirl	-0.0766	-0.0383
hheadsexfemale	-0.0127	-0.0046
edyrtotal	-0.0165	-0.0654
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-0.0599	-0.0299
maternal_age_month	-9e-04	-0.0683
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	2e-04	1e-04
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	0.0559	0.0242
religion_newOther	-0.0625	-0.0157
bednetnumlog	0.1042	0.053
prev_deathyes	0.3372	0.1274
prev_death_fullyes	-	-
malaria_new	-0.3486	-0.037
kidbordlog	0.0233	0.0162
kidbirthmo_sin	9e-04	6e-04
kidbirthmo_cos	-0.0835	-0.0582
geo_cm2011Centre	-	-
geo_cm2011Douala	-0.2219	-0.0542
geo_cm2011Est	-0.2023	-0.0519
geo_cm2011Extrême-Nord	0.1703	0.0628
geo_cm2011Littoral	-	-
geo_cm2011Nord	0.1748	0.0606
geo_cm2011Nord-Ouest	-0.0083	-0.0022
geo_cm2011Ouest	-0.0285	-0.0083
geo_cm2011Sud	-	-
geo_cm2011Sud-Ouest	0.1559	0.0355
geo_cm2011Yaoundé	-	-
wealthp	-	-
wealthp2	-	-

B.6 Cote d'Ivoire

This sample was taken in the year 2011. There are 5211 observations in the dataset, and 4169 observations were used for the training set. Also, 734 observations were removed due to missing values. Finally, there were 421 deaths for the full dataset, and 337 deaths in the training set.

B.6.1 Variables Used

Table 41: Variables included in the model for Cote d'Ivoire

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
wealthp	numeric
wealthp2	numeric

B.6.2 Mortality Breakdown by Wealth Quintile

B.6.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 42: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
1	0.0029452

Table 43: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 44: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 45: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	50	0.0002803

B.6.4 Table of Results

Table 46: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.208	0.395	0.654	0.149	0.018	1.241	0.073	0.919	1.943
KRLS	0.208	0.330	0.638	0.146	0.016	1.205	0.073	0.920	1.947
Elastic Net	0.205	0.321	0.629	0.132	0.016	1.205	0.073	0.920	1.918
Random Forest	0.217	0.353	0.622	0.154	0.024	1.432	0.073	0.921	2.027
Wealth	0.107	0.199	0.529	0.098	0.076	2.132	0.809	0.911	1.000
XGB	0.193	0.324	0.619	0.136	0.015	1.185	0.073	0.919	1.806

Table 47: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.088	0.124	0.105	0.185	0.154
2	0.089	0.118	0.113	0.127	0.084
3	0.115	0.124	0.139	0.098	0.112
4	0.124	0.117	0.129	0.088	0.102
5	0.117	0.102	0.132	0.066	0.105
6	0.139	0.117	0.112	0.083	0.110
7	0.125	0.115	0.120	0.065	0.123
8	0.105	0.100	0.090	0.117	0.098
9	0.049	0.054	0.032	0.059	0.061
10	0.046	0.039	0.024	0.110	0.049

B.6.5 Performance Plots

Below are various plots showing the performance for each model considered in Cote d'Ivoire.

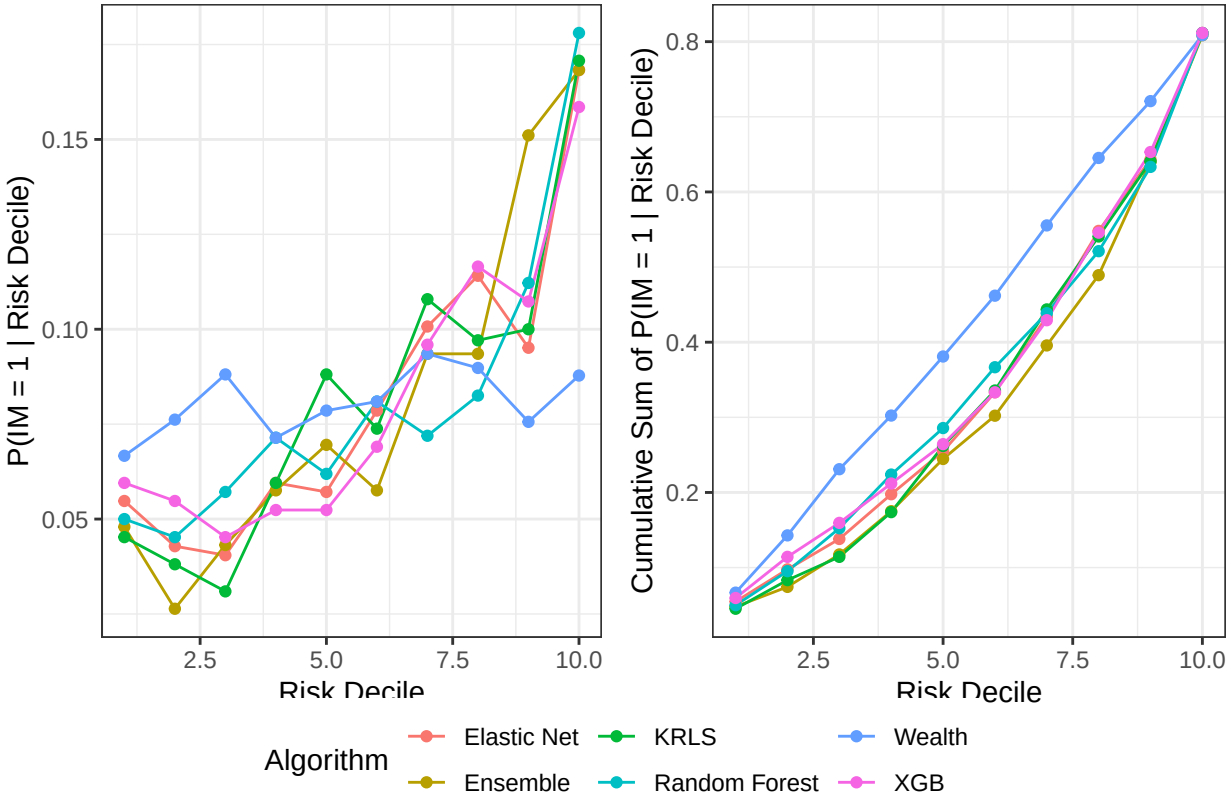

Figure 13: Probability of a mortality given that the observation is in a particular risk decile

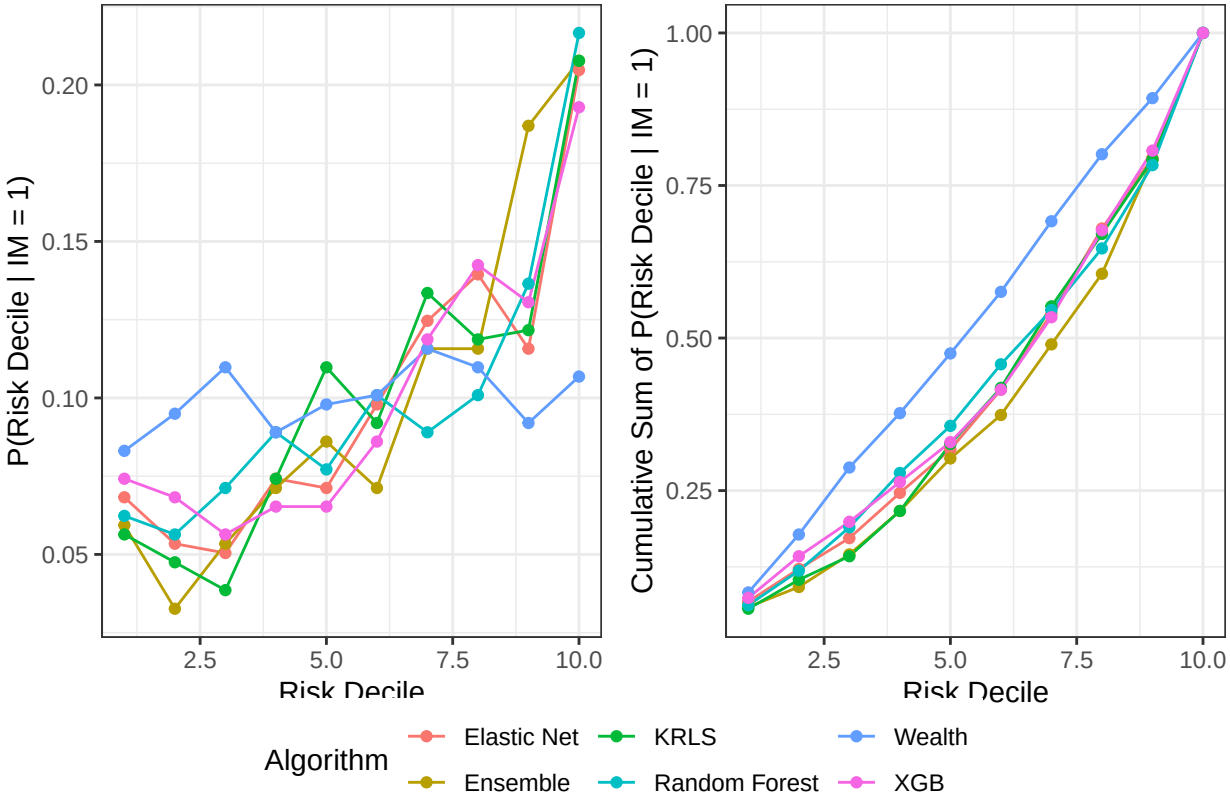

Figure 14: Probability of membership in a particular risk decile given a mortality

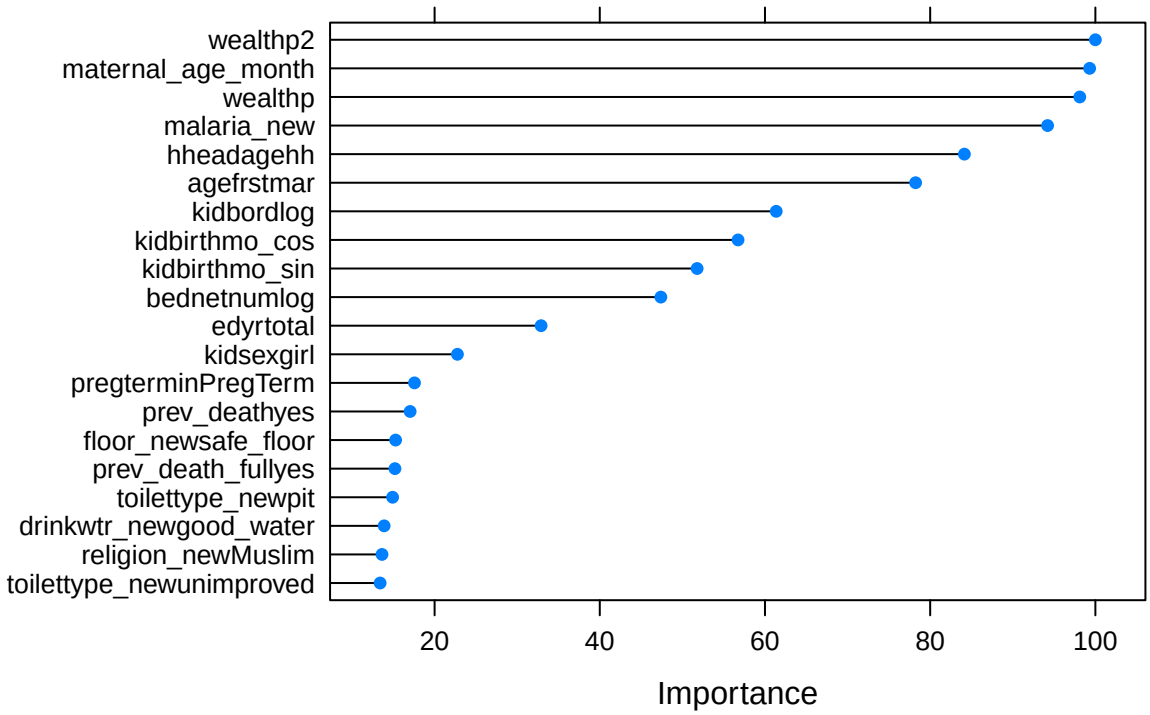

Figure 15: Variable Importance Plot generated by the Random Forest model

B.6.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.6.7 Elastic Net Coefficients

Table 48: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-1.3584	-1.3584
pregterminPregTerm	0.1588	0.0614
agefrstmar	-	-
hheadagehh	-	-
urbanRural	0.3648	0.1701
kidsexgirl	-0.5679	-0.284
hheadsexfemale	-	-
edyrtotal	-	-
drinkwtr_newgood_water	0.0773	0.0334
floor_newsafe_floor	-	-
maternal_age_month	-8e-04	-0.065
cookfuel_newclean_fuel	0.0267	0.0068
toilettype_newpit	0.0909	0.0452
toilettype_newunimproved	-	-
religion_newChristian	-0.0845	-0.04
religion_newHindu	-	-
religion_newMuslim	0.0192	0.0096
religion_newOther	-	-
bednetnumlog	-	-
prev_deathyes	0.5558	0.2327
prev_death_fullyes	0.0971	0.0444
malaria_new	-2.1569	-0.1071
kidbordlog	-	-
kidbirthmo_sin	0.0599	0.0422
kidbirthmo_cos	-	-
wealthp	-	-
wealthp2	-	-

B.7 Ghana

This sample was taken in the year 2014. There are 4203 observations in the dataset, and 3364 observations were used for the training set. Also, 345 observations were removed due to missing values. Finally, there were 189 deaths for the full dataset, and 152 deaths in the training set.

B.7.1 Variables Used

Table 49: Variables included in the model for Ghana

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
wealthp	numeric
wealthp2	numeric

B.7.2 Mortality Breakdown by Wealth Quintile

B.7.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 50: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.6785714	0.0090749

Table 51: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 52: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 53: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	50	0.0086021

B.7.4 Table of Results

Table 54: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.185	0.271	0.539	0.120	0.007	1.161	0.043	0.931	3.127
KRLS	0.165	0.290	0.567	0.065	0.000	1.009	0.043	0.931	2.796
Elastic Net	0.185	0.304	0.519	0.062	0.002	1.048	0.043	0.932	3.127
Random Forest	0.132	0.243	0.556	0.102	0.019	1.530	0.043	0.929	2.232
Wealth	0.059	0.210	0.568	0.046	-0.038	0.889	0.993	0.926	1.000
XGB	0.172	0.285	0.539	0.079	0.006	1.135	0.043	0.931	2.901

Table 55: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.145	0.112	0.133	0.139	0.073
2	0.122	0.149	0.116	0.116	0.155
3	0.164	0.142	0.158	0.127	0.127
4	0.076	0.088	0.079	0.079	0.085
5	0.094	0.052	0.082	0.045	0.061
6	0.133	0.118	0.148	0.091	0.133
7	0.060	0.074	0.080	0.071	0.083
8	0.073	0.079	0.079	0.082	0.052
9	0.052	0.064	0.055	0.073	0.082
10	0.079	0.136	0.067	0.173	0.145

B.7.5 Performance Plots

Below are various plots showing the performance for each model considered in Ghana.

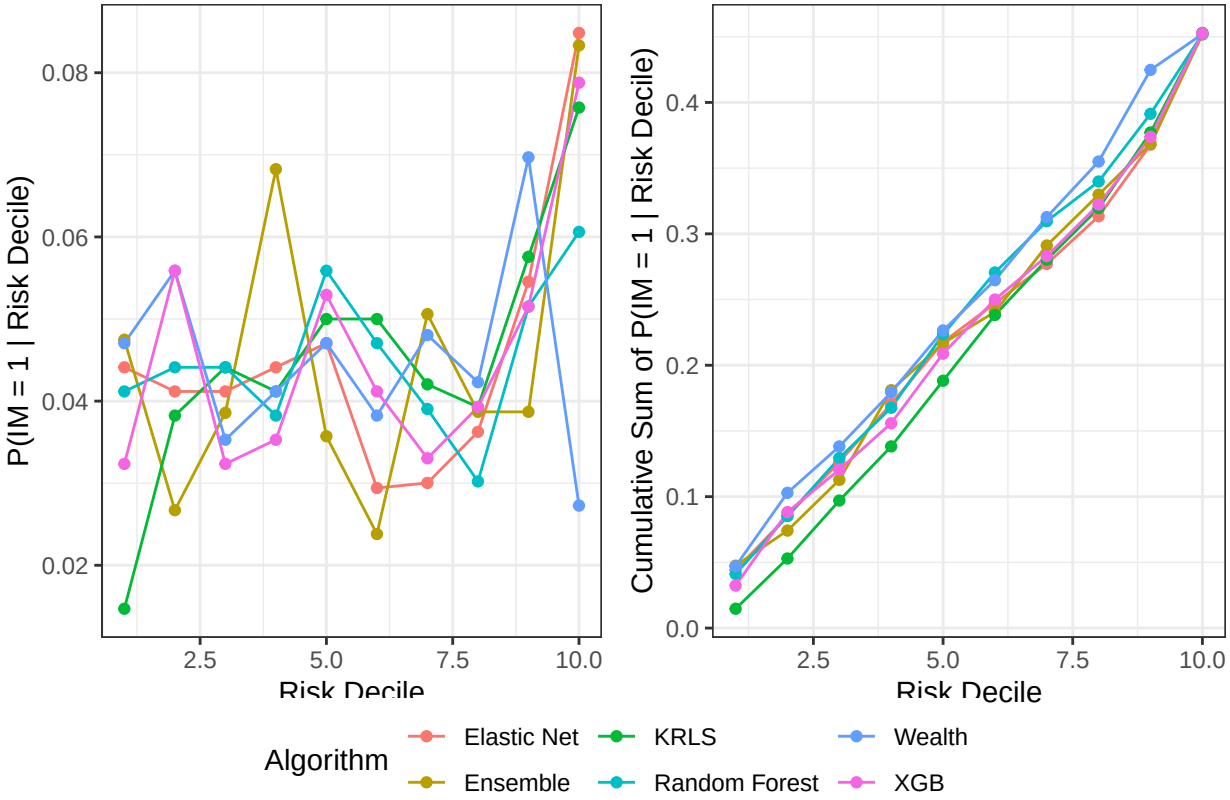

Figure 16: Probability of a mortality given that the observation is in a particular risk decile

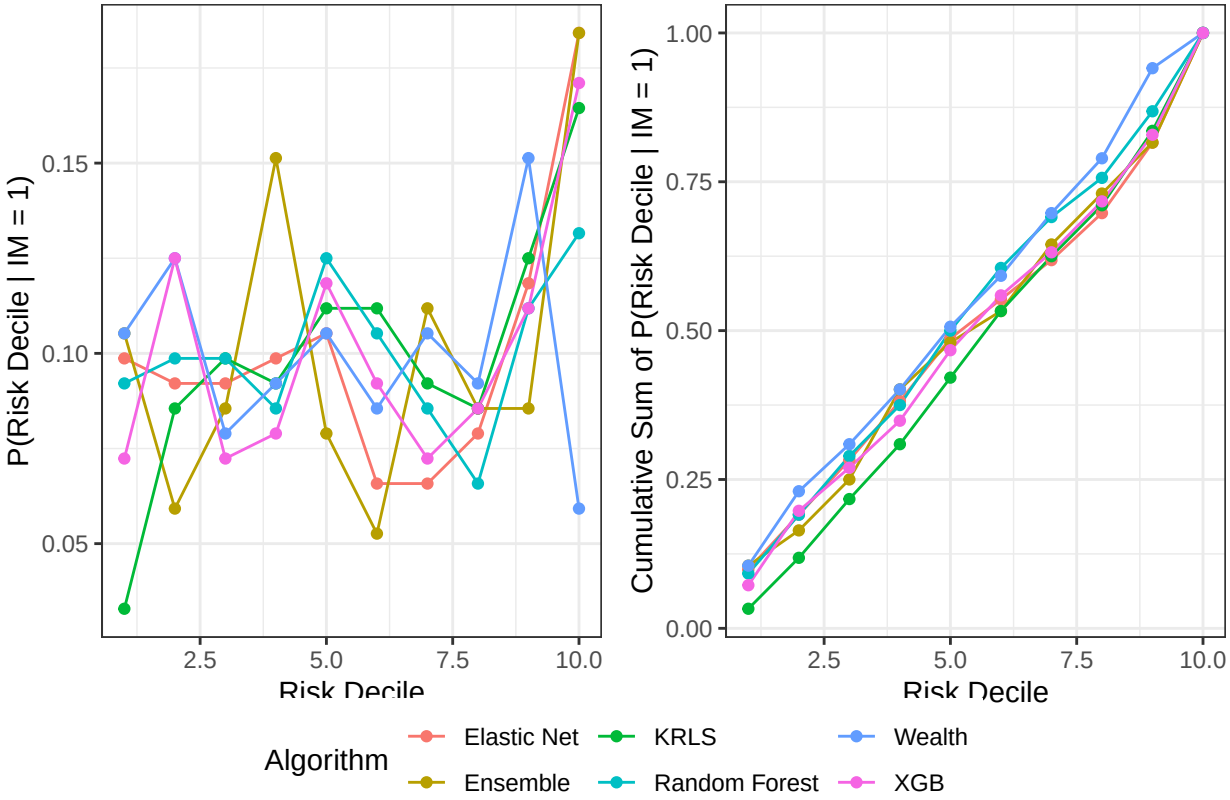

Figure 17: Probability of membership in a particular risk decile given a mortality

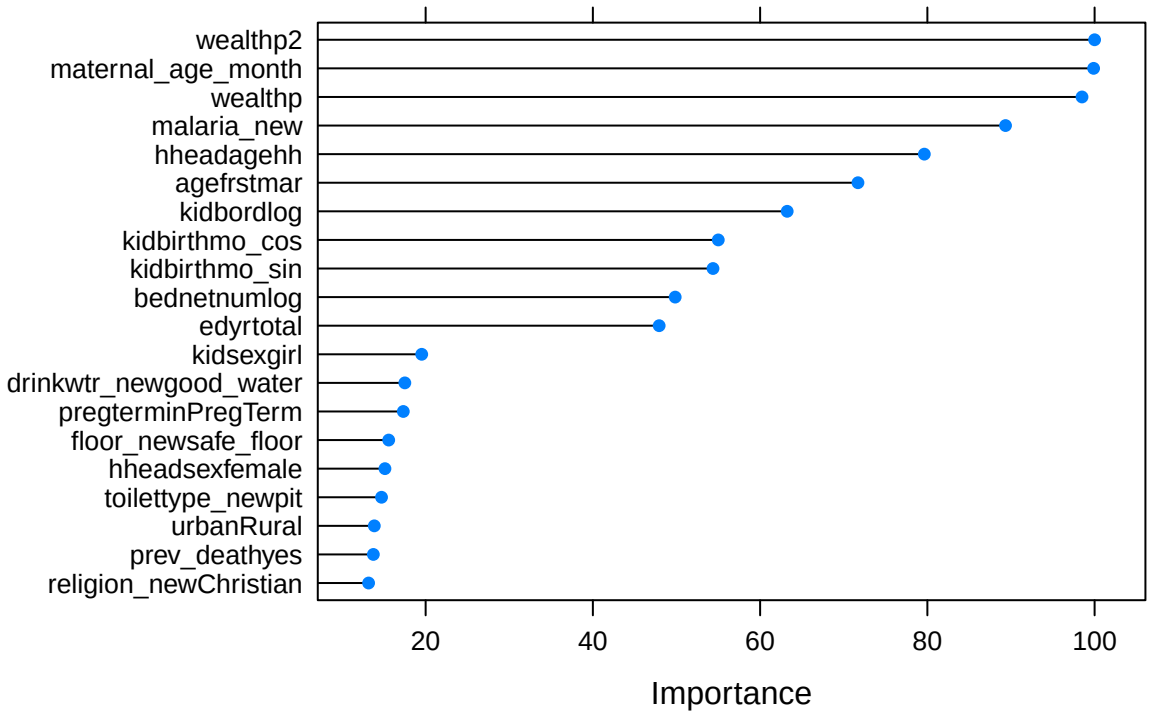

Figure 18: Variable Importance Plot generated by the Random Forest model

B.7.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.7.7 Elastic Net Coefficients

Table 56: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-3.1985	-3.1985
pregterminPregTerm	-	-
agefrstmar	-	1e-04
hheadagehh	-	-
urbanRural	-	-
kidsexgirl	-	-
hheadsexfemale	-	-
edyrtotal	-	-
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	2e-04	0.0162
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-	-
prev_deathyes	0.4296	0.1533
prev_death_fullyes	-	-
malaria_new	-	-
kidbordlog	-	-
kidbirthmo_sin	-	-
kidbirthmo_cos	-	-
wealthp	-	-
wealthp2	-	-

B.8 Guinea

This sample was taken in the year 2012. There are 5223 observations in the dataset, and 4179 observations were used for the training set. Also, 199 observations were removed due to missing values. Finally, there were 396 deaths for the full dataset, and 317 deaths in the training set.

B.8.1 Variables Used

Table 57: Variables included in the model for Guinea

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_gn2012	categorical
wealthp	numeric
wealthp2	numeric

B.8.2 Mortality Breakdown by Wealth Quintile

B.8.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 58: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.4857143	0.0068884

Table 59: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 60: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	0.875

Table 61: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	62	0.000648

B.8.4 Table of Results

Table 62: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.170	0.318	0.606	0.122	0.009	1.134	0.069	0.918	1.001
KRLS	0.157	0.274	0.582	0.116	0.005	1.071	0.070	0.918	0.925
Elastic Net	0.157	0.274	0.576	0.114	0.005	1.070	0.070	0.918	0.926
Random Forest	0.145	0.287	0.582	0.118	0.012	1.301	0.071	0.917	0.853
Wealth	0.170	0.283	0.574	0.108	0.242	2.981	0.896	0.919	1.000
XGB	0.139	0.296	0.599	0.117	0.013	1.171	0.070	0.916	0.819

Table 63: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.298	0.407	0.249	0.320	0.427
2	0.290	0.182	0.283	0.110	0.144
3	0.224	0.144	0.227	0.068	0.132
4	0.110	0.076	0.112	0.078	0.085
5	0.041	0.071	0.061	0.110	0.063
6	0.017	0.049	0.039	0.093	0.054
7	0.012	0.034	0.012	0.067	0.050
8	0.002	0.032	0.012	0.066	0.029
9	0.000	0.015	0.000	0.041	0.012
10	0.000	0.005	0.000	0.044	0.000

B.8.5 Performance Plots

Below are various plots showing the performance for each model considered in Guinea.

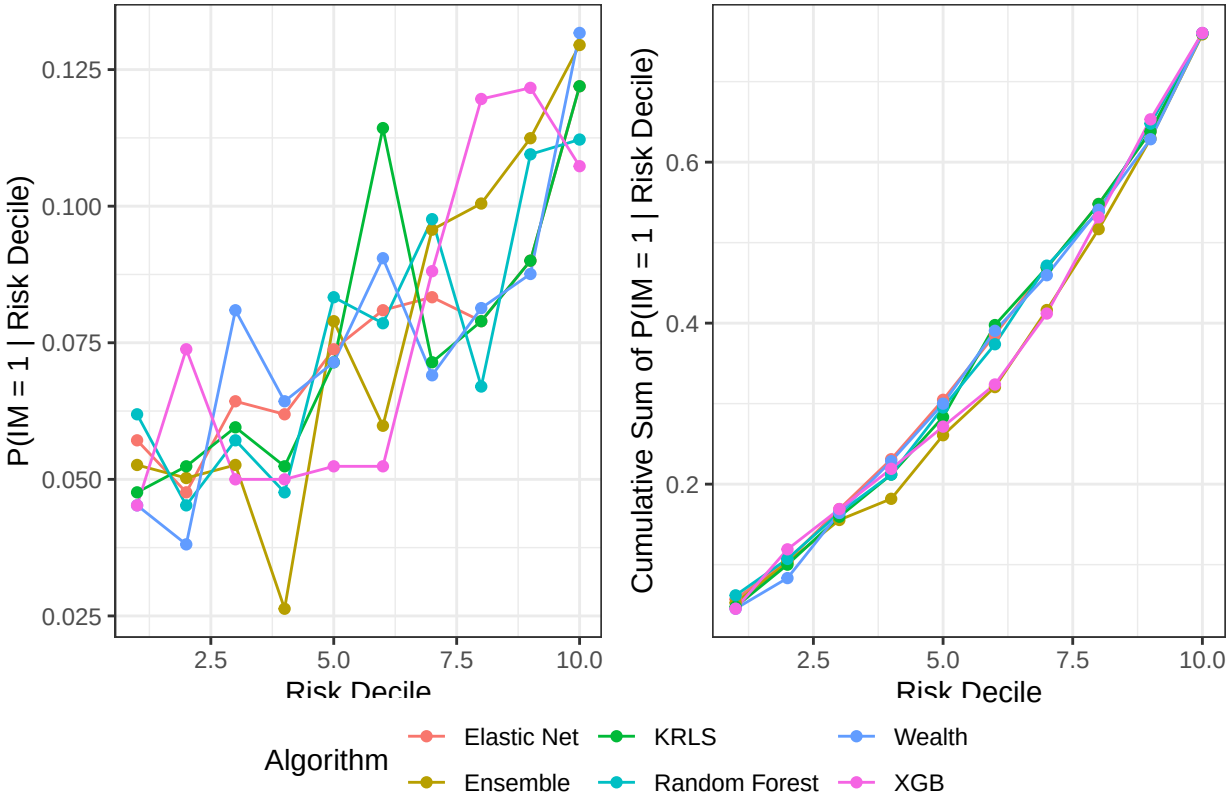

Figure 19: Probability of a mortality given that the observation is in a particular risk decile

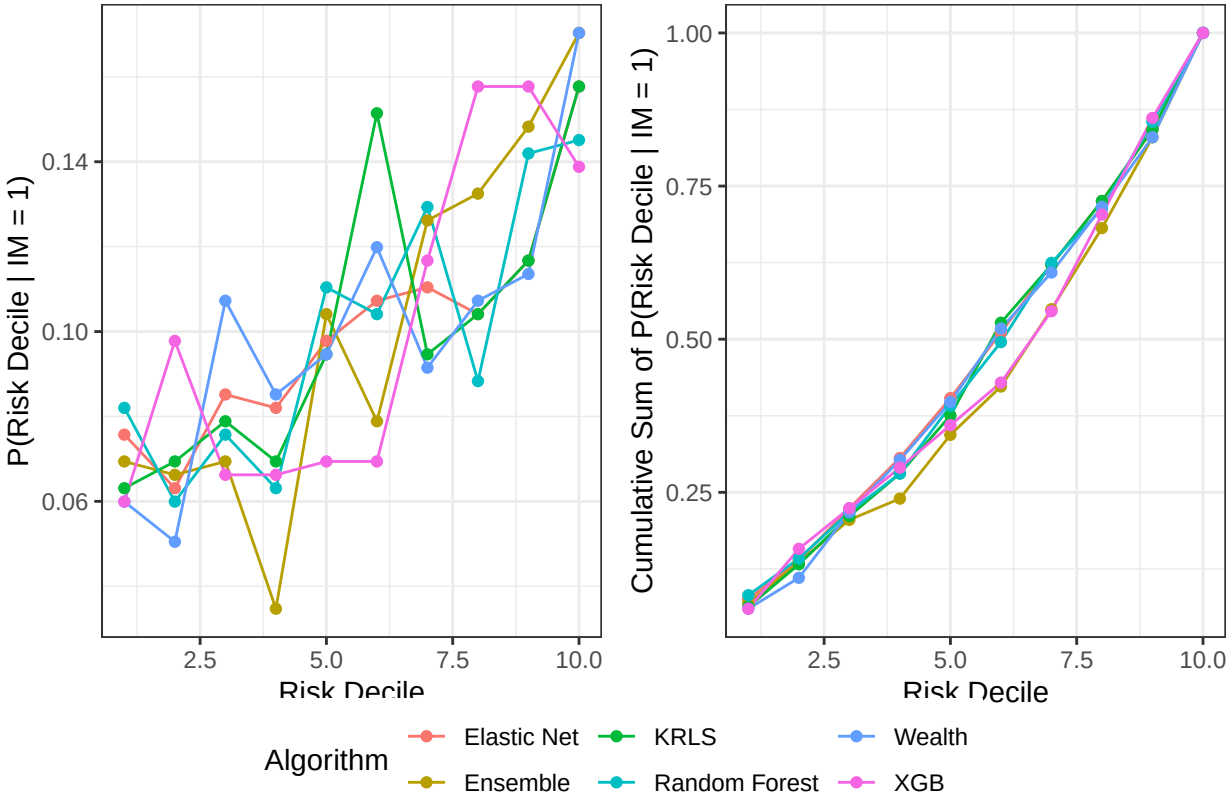

Figure 20: Probability of membership in a particular risk decile given a mortality

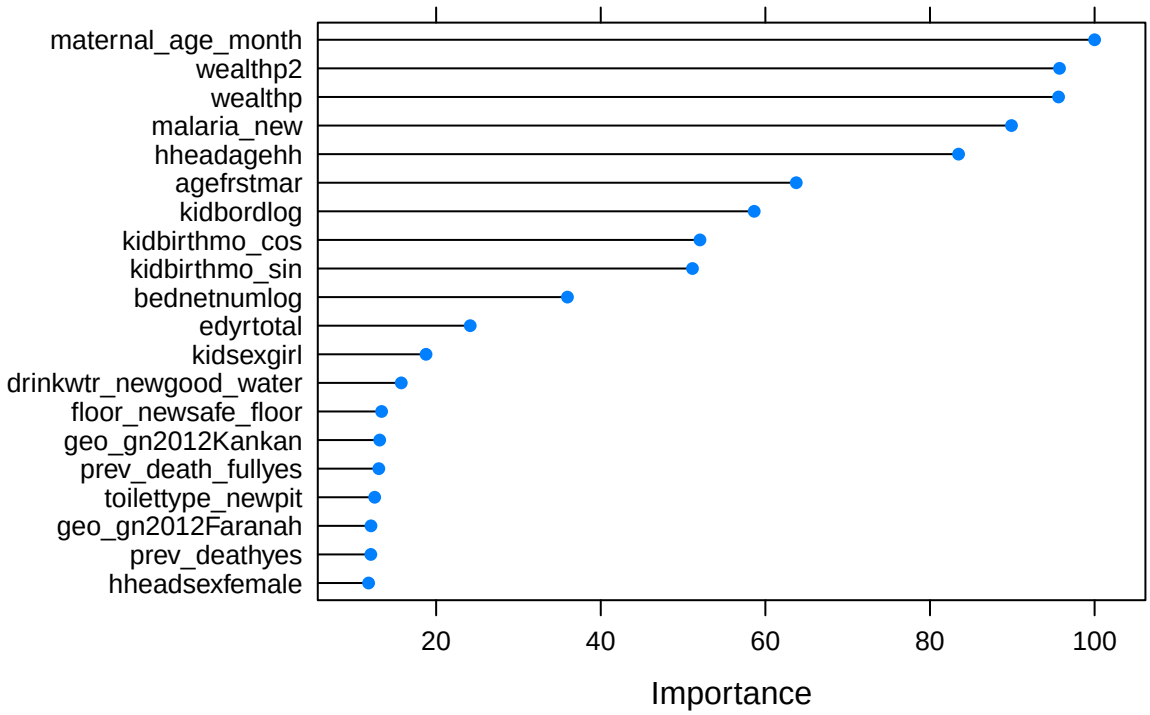

Figure 21: Variable Importance Plot generated by the Random Forest model

B.8.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.8.7 Elastic Net Coefficients

Table 64: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.1617	-2.1617
pregterminPregTerm	-	-
agefrstmar	-0.0066	-0.0233
hheadagehh	-	-
urbanRural	-	-
kidsexgirl	-	-
hheadsexfemale	-	-
edyrtotal	-0.015	-0.0466
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-0.0668	-0.0334
maternal_age_month	-9e-04	-0.0798
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-0.1411	-0.0785
prev_deathyes	0.0065	0.0027
prev_death_fullyes	0.043	0.0196
malaria_new	0.9926	0.1039
kidbordlog	-	-
kidbirthmo_sin	-0.0033	-0.0023
kidbirthmo_cos	0.0484	0.0332
geo_gn2012Conakry	-	-
geo_gn2012Faranah	-	-
geo_gn2012Kankan	0.133	0.0504
geo_gn2012Kindia	-0.1895	-0.0613
geo_gn2012Labe	-	-
geo_gn2012Mamou	-	-
geo_gn2012N'Zerekore	-0.1833	-0.0594
wealthp	-0.4189	-0.1214
wealthp2	-	-

B.9 Kenya

This sample was taken in the year 2014. There are 15110 observations in the dataset, and 12089 observations were used for the training set. Also, 1447 observations were removed due to missing values. Finally, there were 589 deaths for the full dataset, and 472 deaths in the training set.

B.9.1 Variables Used

Table 65: Variables included in the model for Kenya

Variable	**Type**
mortality.under12m	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
wealthp	numeric
wealthp2	numeric

B.9.2 Mortality Breakdown by Wealth Quintile

B.9.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 66: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.2285714	0.0082947

Table 67: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 68: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 69: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	48	0.004079

B.9.4 Table of Results

Table 70: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.163	0.301	0.573	0.065	0.003	1.088	0.037	0.932	2.143
KRLS	0.150	0.250	0.536	0.048	0.000	1.005	0.038	0.932	1.975
Elastic Net	0.148	0.271	0.524	0.051	0.001	1.025	0.037	0.932	1.948
Random Forest	0.168	0.282	0.564	0.066	0.009	1.408	0.038	0.933	2.199
Wealth	0.076	0.167	0.542	0.039	-0.028	0.882	1.081	0.929	1.000
XGB	0.127	0.273	0.558	0.047	0.002	1.046	0.038	0.931	1.668

Table 71: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.092	0.059	0.053	0.076	0.022
2	0.105	0.078	0.079	0.092	0.044
3	0.117	0.127	0.112	0.116	0.127
4	0.123	0.108	0.133	0.092	0.139
5	0.119	0.115	0.130	0.095	0.127
6	0.125	0.096	0.131	0.076	0.127
7	0.099	0.092	0.115	0.086	0.105
8	0.093	0.111	0.107	0.115	0.090
9	0.074	0.082	0.085	0.087	0.073
10	0.052	0.137	0.054	0.164	0.143

B.9.5 Performance Plots

Below are various plots showing the performance for each model considered in Kenya.

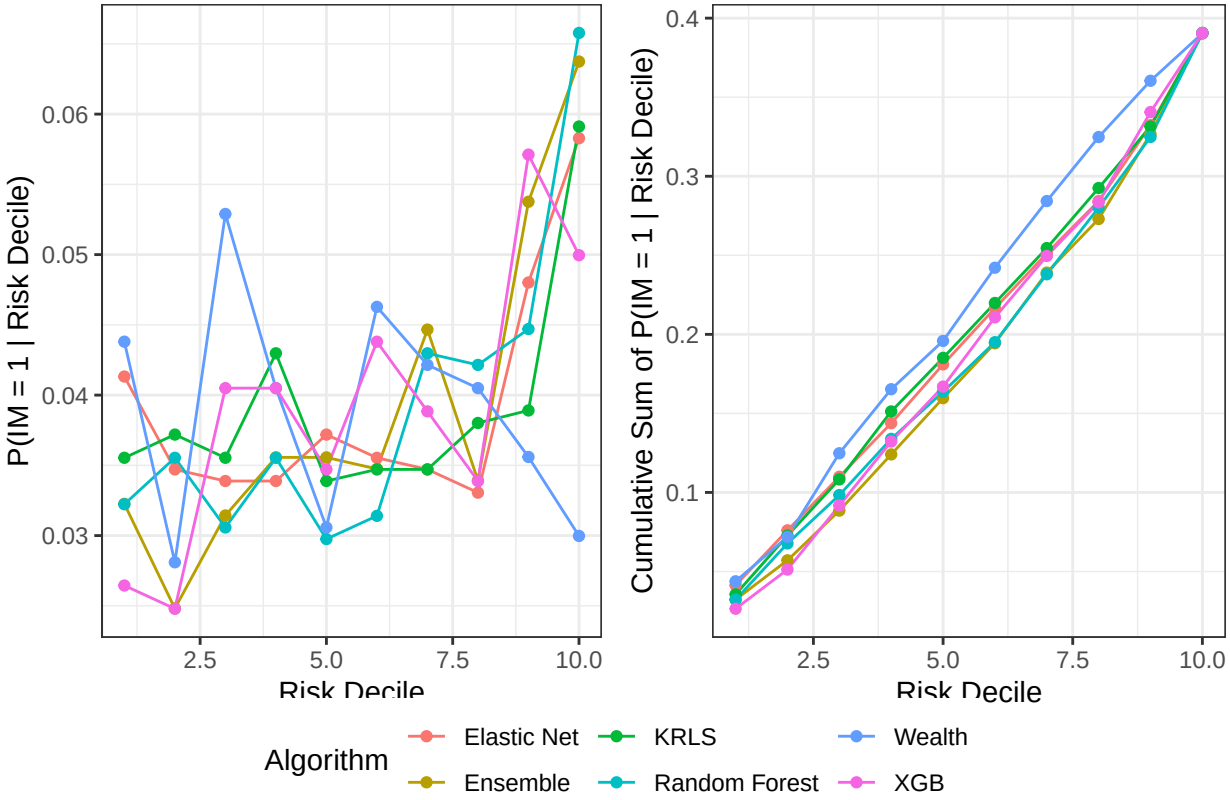

Figure 22: Probability of a mortality given that the observation is in a particular risk decile

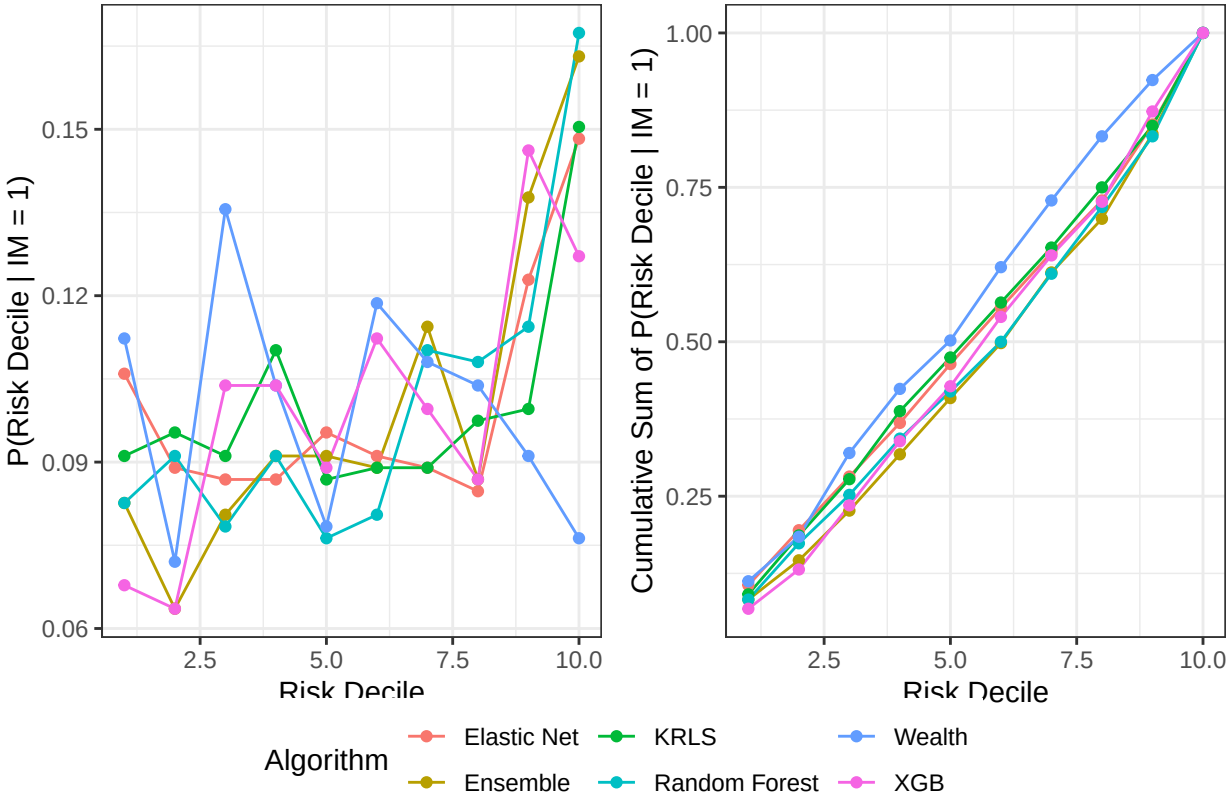

Figure 23: Probability of membership in a particular risk decile given a mortality

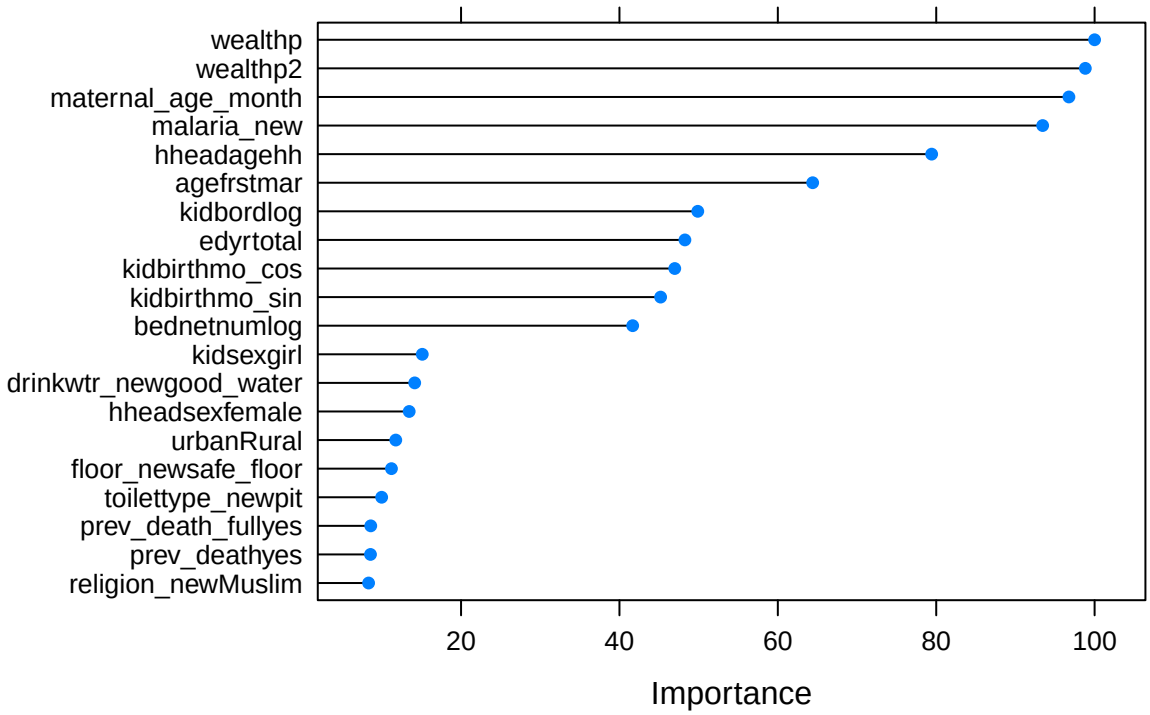

Figure 24: Variable Importance Plot generated by the Random Forest model

B.9.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.9.7 Elastic Net Coefficients

Table 72: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-3.2867	-3.2867
agefrstmar	-	-
hheadagehh	0.0011	0.0132
urbanRural	-	-
kidsexgirl	-0.0811	-0.0406
hheadsexfemale	-	-
edyrtotal	-	-
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	-	-
cookfuel_newclean_fuel	-	-
toilettype_newpit	0.0494	0.0234
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-0.0235	-0.0132
prev_deathyes	0.1707	0.0544
prev_death_fullyes	0.2008	0.0704
malaria_new	0.0167	0.0019
kidbordlog	-	-
kidbirthmo_sin	0.0688	0.0488
kidbirthmo_cos	-	-
wealthp	0.0063	0.0018
wealthp2	-	-

B.10 Lesotho

This sample was taken in the year 2014. There are 2127 observations in the dataset, and 1702 observations were used for the training set. Also, 233 observations were removed due to missing values. Finally, there were 145 deaths for the full dataset, and 116 deaths in the training set.

B.10.1 Variables Used

Table 73: Variables included in the model for Lesotho

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_ls2014	categorical
wealthp	numeric
wealthp2	numeric

B.10.2 Mortality Breakdown by Wealth Quintile

B.10.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 74: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.6785714	0.019543

Table 75: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 76: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	1

Table 77: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	64	0.001915

B.10.4 Table of Results

Table 78: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.138	0.267	0.586	0.112	0.005	1.076	0.063	0.918	2.022
KRLS	0.145	0.249	0.548	0.089	0.001	1.017	0.063	0.920	2.133
Elastic Net	0.094	0.214	0.575	0.083	0.001	1.011	0.063	0.916	1.378
Random Forest	0.162	0.317	0.582	0.128	0.013	1.344	0.064	0.921	2.378
Wealth	0.068	0.137	0.583	0.079	-0.026	0.859	0.993	0.914	1.000
XGB	0.077	0.181	0.522	0.087	0.003	1.047	0.064	0.915	1.133

Table 79: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.268	0.054	0.060	0.089	0.054
2	0.124	0.024	0.053	0.065	0.000
3	0.214	0.137	0.143	0.125	0.113
4	0.113	0.065	0.071	0.060	0.024
5	0.071	0.060	0.107	0.071	0.036
6	0.048	0.065	0.125	0.083	0.042
7	0.076	0.241	0.176	0.147	0.282
8	0.054	0.185	0.131	0.161	0.238
9	0.012	0.101	0.071	0.089	0.137
10	0.018	0.077	0.060	0.107	0.071

B.10.5 Performance Plots

Below are various plots showing the performance for each model considered in Lesotho.

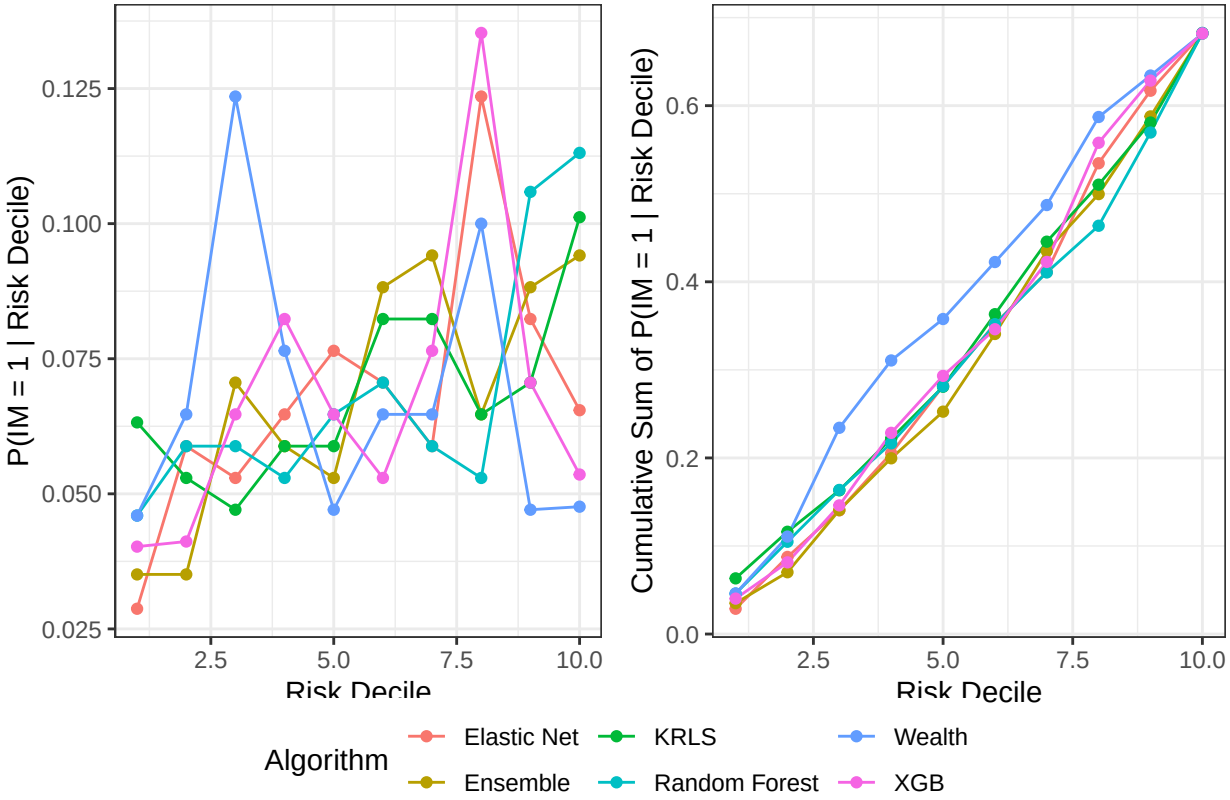

Figure 25: Probability of a mortality given that the observation is in a particular risk decile

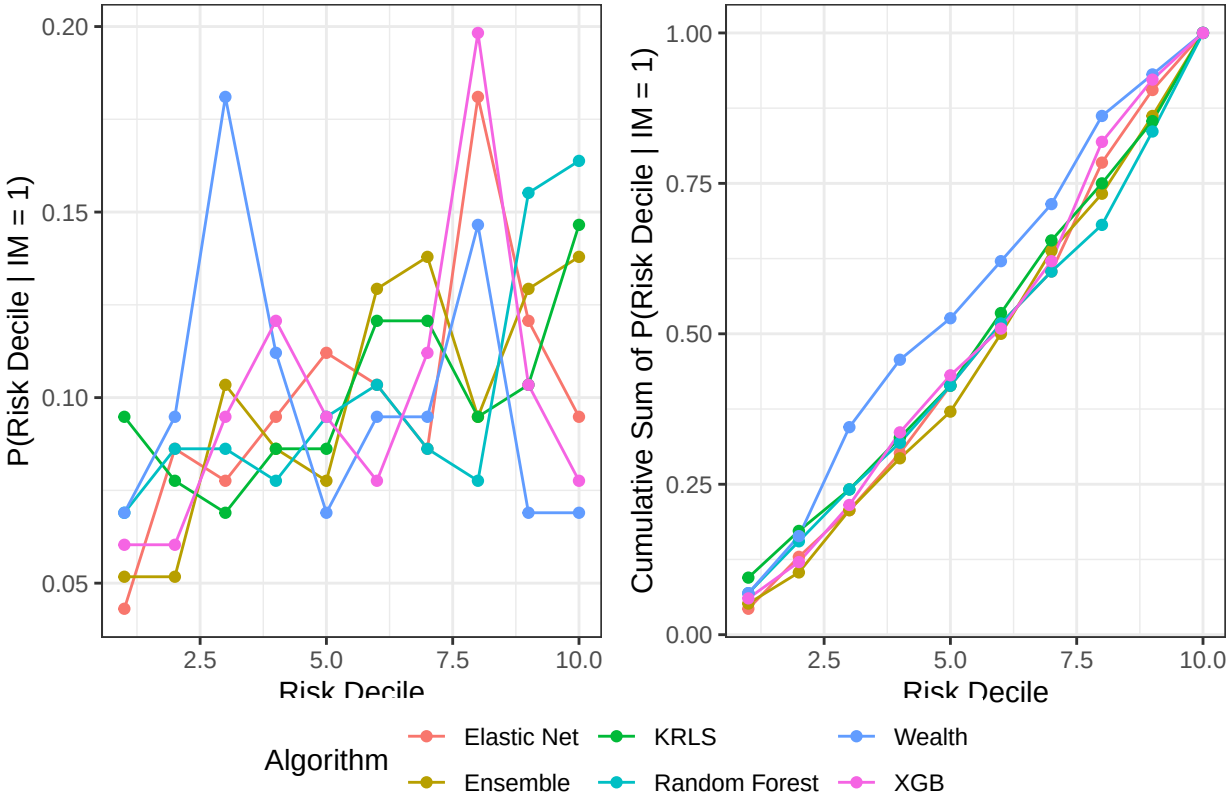

Figure 26: Probability of membership in a particular risk decile given a mortality

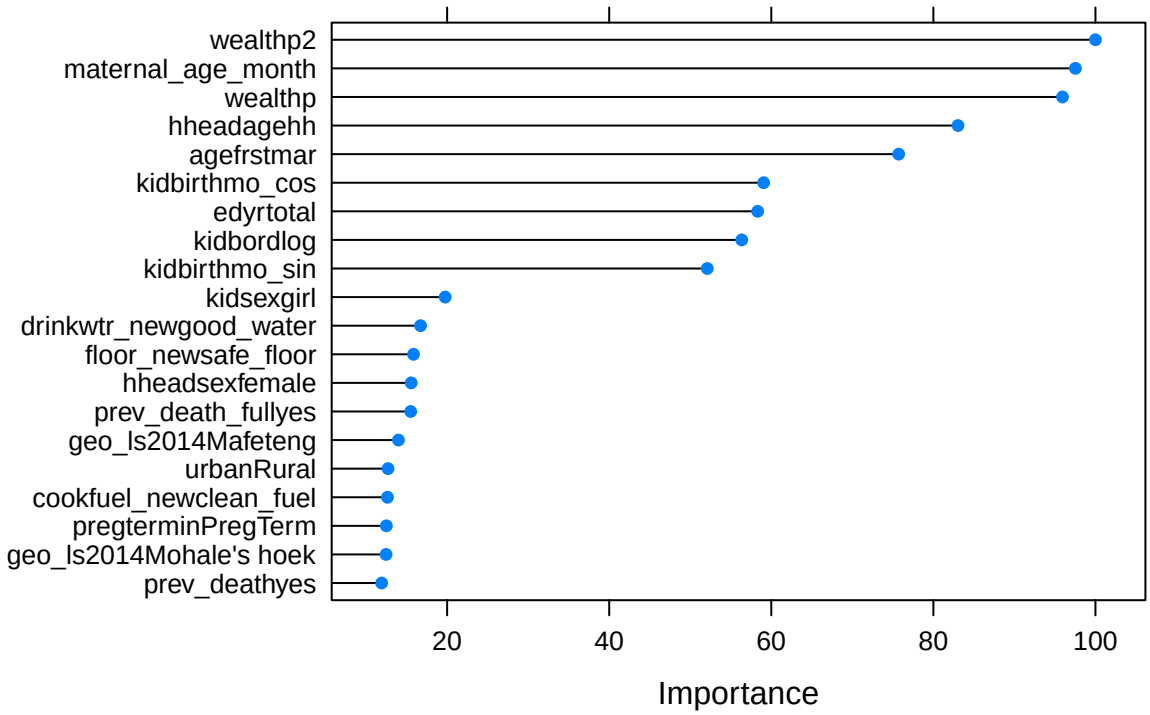

Figure 27: Variable Importance Plot generated by the Random Forest model

B.10.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.10.7 Elastic Net Coefficients

Table 80: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.439	-2.439
pregterminPregTerm	-	-
agefrstmar	-	-
hheadagehh	-	-
urbanRural	-	-
kidsexgirl	-	-
hheadsexfemale	-	-
edyrtotal	-0.0208	-0.06
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	-	-
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
prev_deathyes	-	-
prev_death_fullyes	-	-
kidbordlog	-	-
kidbirthmo_sin	-	-
kidbirthmo_cos	-	-
geo_ls2014Leribe	-	-
geo_ls2014Berea	-	-
geo_ls2014Maseru	-	-
geo_ls2014Mafeteng	-	-
geo_ls2014Mohale's hoek	-	-
geo_ls2014Quthing	-	-
geo_ls2014Qacha's-nek	-	-
geo_ls2014Mokhotlong	-	-
geo_ls2014Thaba Tseka	-	-
wealthp	-	-
wealthp2	-	-

B.11 Madagascar

This sample was taken in the year 2008. There are 9229 observations in the dataset, and 7384 observations were used for the training set. Also, 428 observations were removed due to missing values. Finally, there were 459 deaths for the full dataset, and 368 deaths in the training set.

B.11.1 Variables Used

Table 81: Variables included in the model for Madagascar

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_mg2008	categorical
wealthp	numeric
wealthp2	numeric

B.11.2 Mortality Breakdown by Wealth Quintile

B.11.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 82: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.8714286	0.0026119

Table 83: Optimal Parameters for the Random Forest algorithm

mtry
13

Table 84: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	0.875

Table 85: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	90	0.000118

B.11.4 Table of Results

Table 86: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.274	0.424	0.656	0.148	0.024	1.473	0.046	0.934	2.652
KRLS	0.250	0.391	0.642	0.100	0.014	1.288	0.047	0.933	2.417
Elastic Net	0.242	0.378	0.629	0.097	0.013	1.270	0.047	0.933	2.339
Random Forest	0.236	0.361	0.635	0.141	0.052	1.933	0.047	0.932	2.285
Wealth	0.103	0.223	0.549	0.053	0.126	1.532	0.795	0.925	1.000
XGB	0.245	0.391	0.653	0.098	0.016	1.338	0.047	0.933	2.364

Table 87: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.132	0.134	0.126	0.138	0.107
2	0.146	0.152	0.156	0.141	0.157
3	0.129	0.129	0.136	0.101	0.112
4	0.100	0.115	0.108	0.099	0.126
5	0.123	0.123	0.136	0.092	0.129
6	0.111	0.100	0.115	0.085	0.112
7	0.103	0.092	0.093	0.107	0.106
8	0.079	0.073	0.077	0.086	0.088
9	0.049	0.049	0.041	0.074	0.045
10	0.025	0.041	0.010	0.074	0.015

B.11.5 Performance Plots

Below are various plots showing the performance for each model considered in Madagascar.

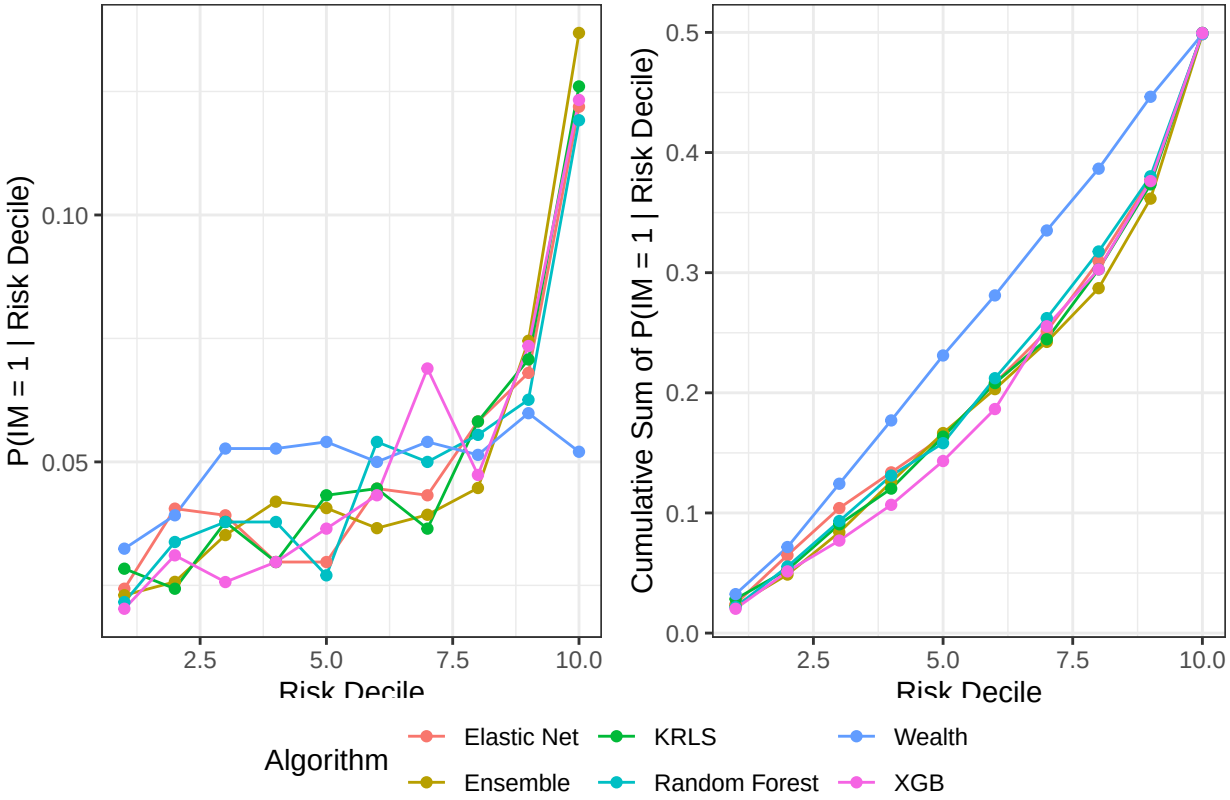

Figure 28: Probability of a mortality given that the observation is in a particular risk decile

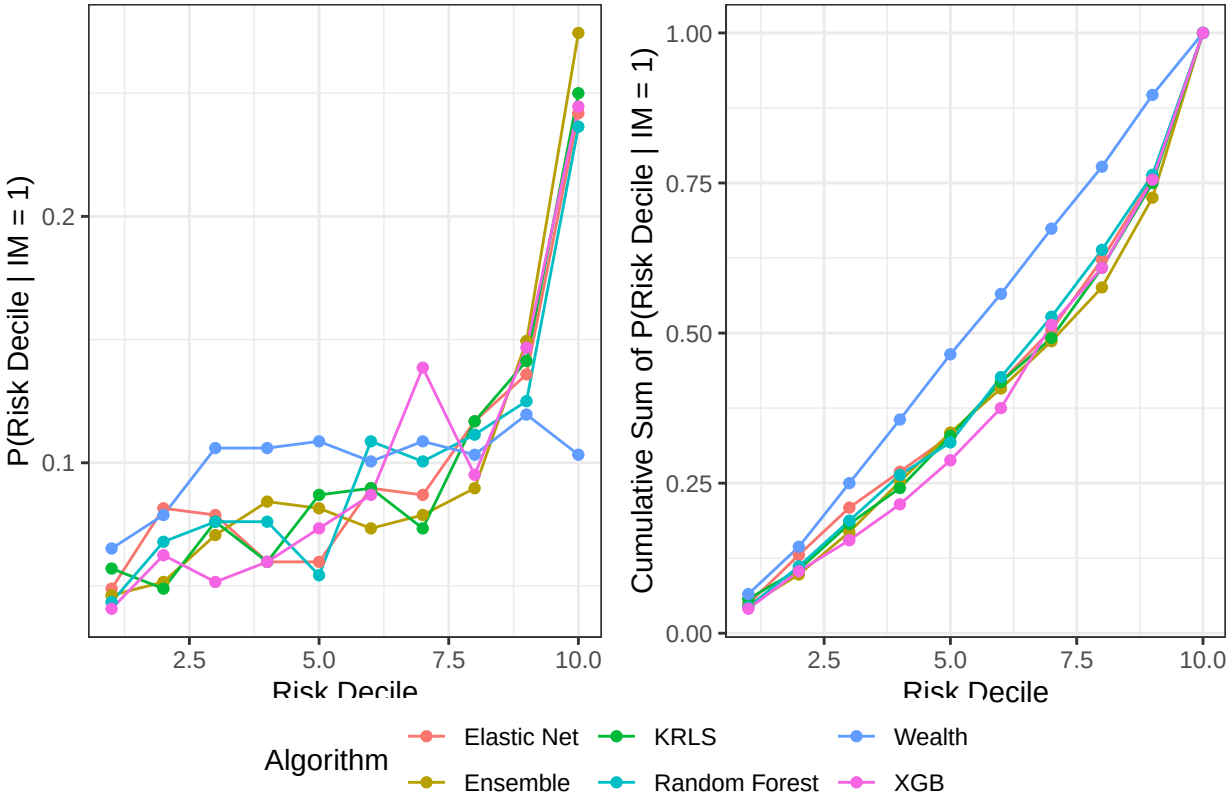

Figure 29: Probability of membership in a particular risk decile given a mortality

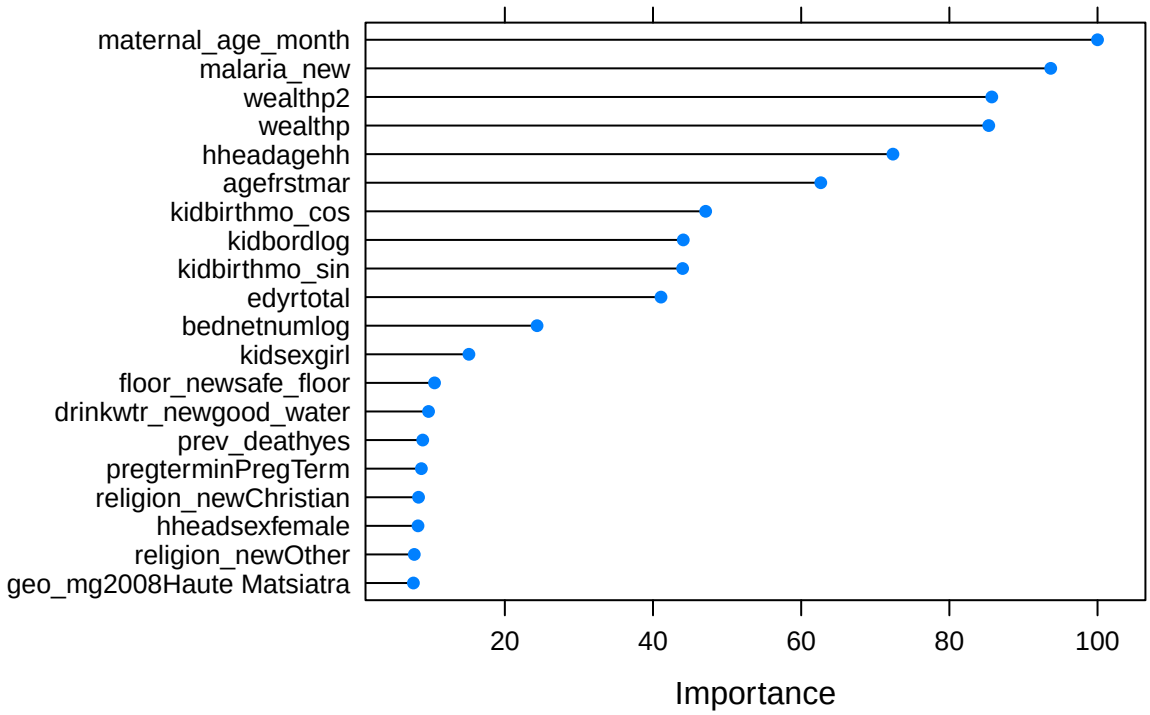

Figure 30: Variable Importance Plot generated by the Random Forest model

B.11.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.11.7 Elastic Net Coefficients

Table 88: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-3.333	-3.333
pregterminPregTerm	-	-
agefrstmar	0.0054	0.021
hheadagehh	3e-04	0.0042
urbanRural	0.0413	0.0157
kidsexgirl	-0.0498	-0.0249
hheadsexfemale	-	-
edyrtotal	-0.0014	-0.0047
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-0.0063	-0.003
maternal_age_month	-	-
toilettype_newpit	-	-
toilettype_newunimproved	0.2149	0.1065
religion_newChristian	0.1559	0.0744
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-0.2581	-0.1232
prev_deathyes	0.9229	0.3388
prev_death_fullyes	-	-
malaria_new	-	-
kidbordlog	-0.0541	-0.039
kidbirthmo_sin	0.1048	0.0738
kidbirthmo_cos	-0.0615	-0.0436
geo_mg2008Vakinankaratra	-	-
geo_mg2008Itasy	-	-
geo_mg2008Bongolava	-	-
geo_mg2008Haute Matsiatra	0.3168	0.071
geo_mg2008Anamoroni'i Mania	-	-
geo_mg2008Vatovavy Fitovinany	0.3908	0.088
geo_mg2008Ihorombe	-	-
geo_mg2008Atsimo Atsinanana	-	-
geo_mg2008Atsinanana	-	-
geo_mg2008Analanjirofo	-	-
geo_mg2008Alaotra Mangoro	-	-
geo_mg2008Boeny	0.3558	0.0689
geo_mg2008Sofia	-0.1246	-0.0253
geo_mg2008Betsiboka	-	-
geo_mg2008Melaky	-0.4814	-0.085

Table 88: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation. *(continued)*

	Coefficients	**Standardized Coefficients**
geo_mg2008Atsimo Andrefana	-0.0529	-0.0113
geo_mg2008Androy	-0.0196	-0.0043
geo_mg2008Anosy	0.081	0.0159
geo_mg2008Menabe	-	-
geo_mg2008Diana	-0.1846	-0.0286
geo_mg2008Sava	-	-
wealthp	-	-
wealthp2	-	-

B.12 Malawi

This sample was taken in the year 2016. There are 13231 observations in the dataset, and 10586 observations were used for the training set. Also, 402 observations were removed due to missing values. Finally, there were 563 deaths for the full dataset, and 451 deaths in the training set.

B.12.1 Variables Used

Table 89: Variables included in the model for Malawi

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_mw2016	categorical
wealthp	numeric
wealthp2	numeric

B.12.2 Mortality Breakdown by Wealth Quintile

B.12.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 90: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1642857	0.0052521

Table 91: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 92: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	0.5

Table 93: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	52	6.05e-05

B.12.4 Table of Results

Table 94: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.211	0.350	0.619	0.097	0.009	1.238	0.040	0.933	1.937
KRLS	0.202	0.321	0.594	0.071	0.006	1.153	0.041	0.933	1.857
Elastic Net	0.164	0.301	0.582	0.061	0.004	1.086	0.041	0.931	1.508
Random Forest	0.204	0.310	0.600	0.115	0.017	1.910	0.040	0.933	1.878
Wealth	0.109	0.215	0.533	0.047	0.063	1.456	0.824	0.929	1.000
XGB	0.184	0.319	0.594	0.068	0.007	1.175	0.041	0.932	1.692

Table 95: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.280	0.218	0.227	0.195	0.121
2	0.150	0.142	0.129	0.104	0.152
3	0.111	0.135	0.130	0.096	0.177
4	0.114	0.129	0.128	0.108	0.151
5	0.070	0.073	0.092	0.059	0.079
6	0.089	0.062	0.091	0.064	0.086
7	0.060	0.061	0.060	0.063	0.062
8	0.050	0.064	0.069	0.082	0.062
9	0.042	0.056	0.042	0.092	0.050
10	0.033	0.066	0.030	0.135	0.057

B.12.5 Performance Plots

Below are various plots showing the performance for each model considered in Malawi.

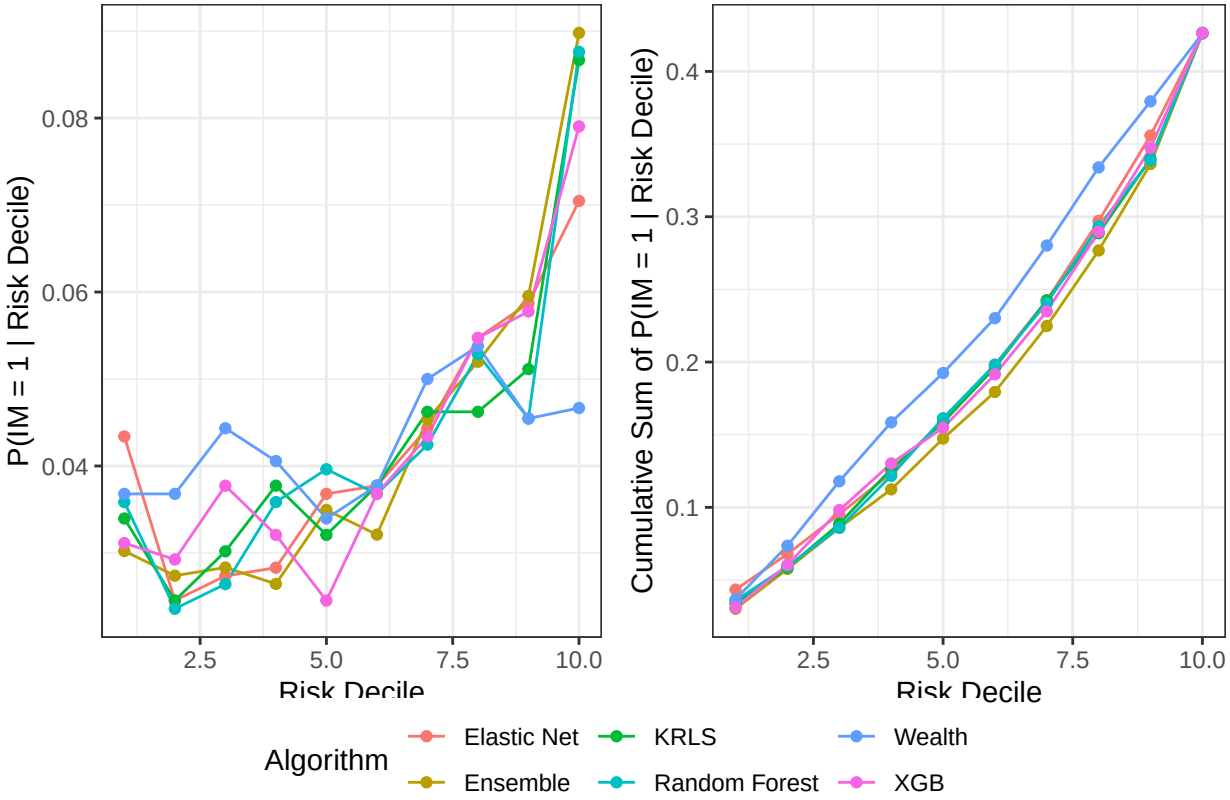

Figure 31: Probability of a mortality given that the observation is in a particular risk decile

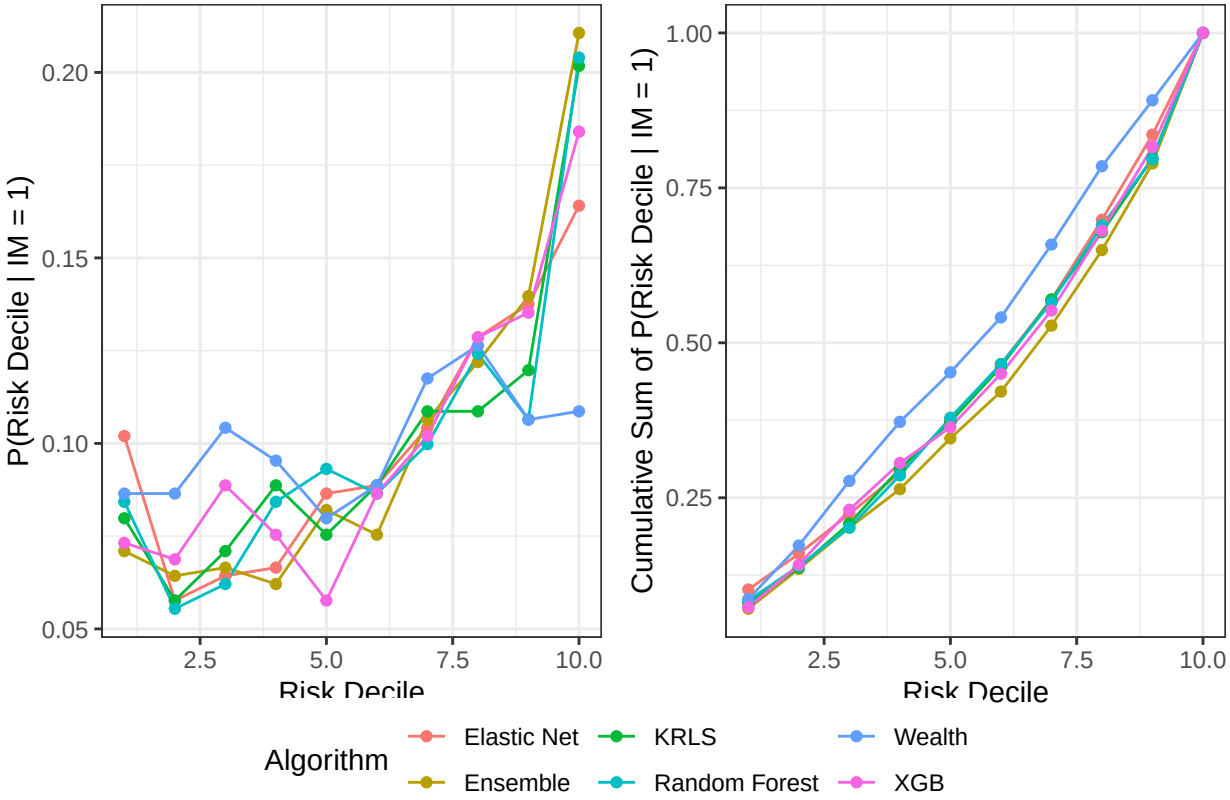

Figure 32: Probability of membership in a particular risk decile given a mortality

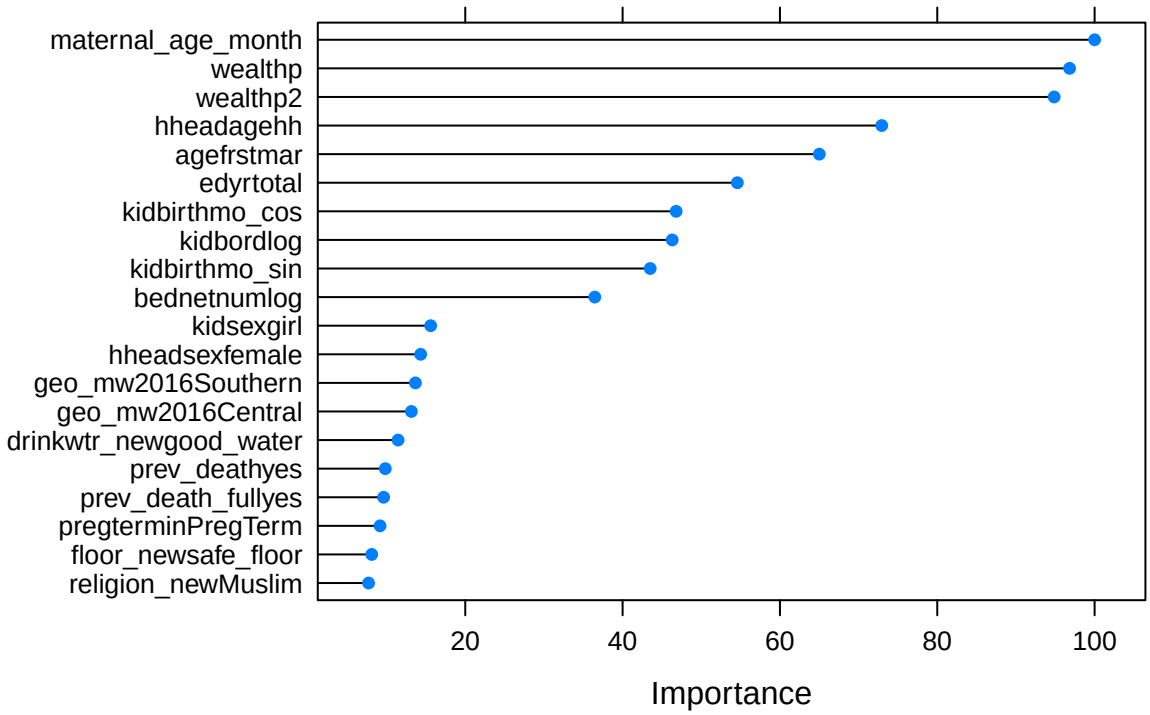

Figure 33: Variable Importance Plot generated by the Random Forest model

B.12.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.12.7 Elastic Net Coefficients

Table 96: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.8538	-2.8538
pregterminPregTerm	-0.0339	-0.0103
agefrstmar	0.0058	0.0194
hheadagehh	-8e-04	-0.0088
urbanRural	0.0638	0.023
kidsexgirl	-0.1948	-0.0974
hheadsexfemale	0.0853	0.0363
edyrtotal	-0.0183	-0.0653
drinkwtr_newgood_water	0.0281	0.0109
floor_newsafe_floor	-2e-04	-1e-04
maternal_age_month	-	-
cookfuel_newclean_fuel	-	-
toilettype_newpit	-0.2427	-0.0694
toilettype_newunimproved	0.2467	0.0623
religion_newChristian	0.0899	0.0311
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-0.4883	-0.0393
bednetnumlog	-	-
prev_deathyes	0.3763	0.1294
prev_death_fullyes	0.2103	0.0827
kidbordlog	-0.2498	-0.1686
kidbirthmo_sin	-0.0014	-0.001
kidbirthmo_cos	0.0452	0.032
geo_mw2016Central	0.0838	0.04
geo_mw2016Southern	-	-
wealthp	-	-
wealthp2	-0.0035	-0.001

B.13 Mali

This sample was taken in the year 2012. There are 7913 observations in the dataset, and 6331 observations were used for the training set. Also, 130 observations were removed due to missing values. Finally, there were 475 deaths for the full dataset, and 380 deaths in the training set.

B.13.1 Variables Used

Table 97: Variables included in the model for Mali

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_ml2012	categorical
wealthp	numeric
wealthp2	numeric

B.13.2 Mortality Breakdown by Wealth Quintile

B.13.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 98: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.0091994

Table 99: Optimal Parameters for the Random Forest algorithm

mtry
8

Table 100: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 101: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	56	0.0001062

B.13.4 Table of Results

Table 102: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.245	0.371	0.666	0.147	0.025	1.410	0.055	0.929	2.114
KRLS	0.213	0.379	0.657	0.110	0.018	1.299	0.055	0.927	1.841
Elastic Net	0.203	0.355	0.646	0.105	0.015	1.262	0.056	0.927	1.750
Random Forest	0.221	0.347	0.634	0.142	0.050	1.742	0.057	0.928	1.909
Wealth	0.116	0.216	0.555	0.068	0.181	3.474	0.833	0.921	1.000
XGB	0.208	0.358	0.654	0.112	0.017	1.291	0.055	0.927	1.795

Table 103: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.119	0.132	0.111	0.148	0.108
2	0.136	0.109	0.125	0.104	0.111
3	0.111	0.081	0.090	0.067	0.100
4	0.124	0.146	0.138	0.132	0.154
5	0.130	0.148	0.146	0.133	0.156
6	0.152	0.171	0.154	0.151	0.187
7	0.114	0.122	0.118	0.112	0.141
8	0.059	0.048	0.067	0.049	0.019
9	0.044	0.030	0.041	0.038	0.017
10	0.010	0.017	0.008	0.065	0.006

B.13.5 Performance Plots

Below are various plots showing the performance for each model considered in Mali.

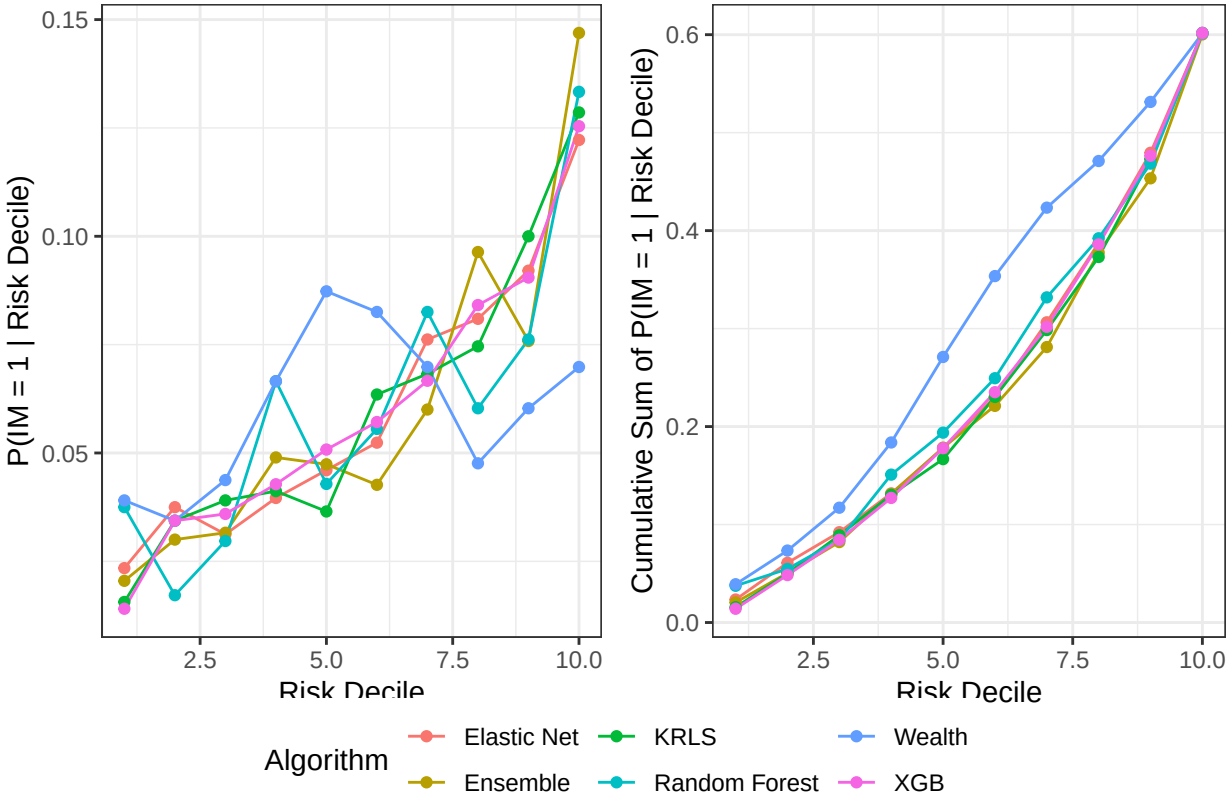

Figure 34: Probability of a mortality given that the observation is in a particular risk decile

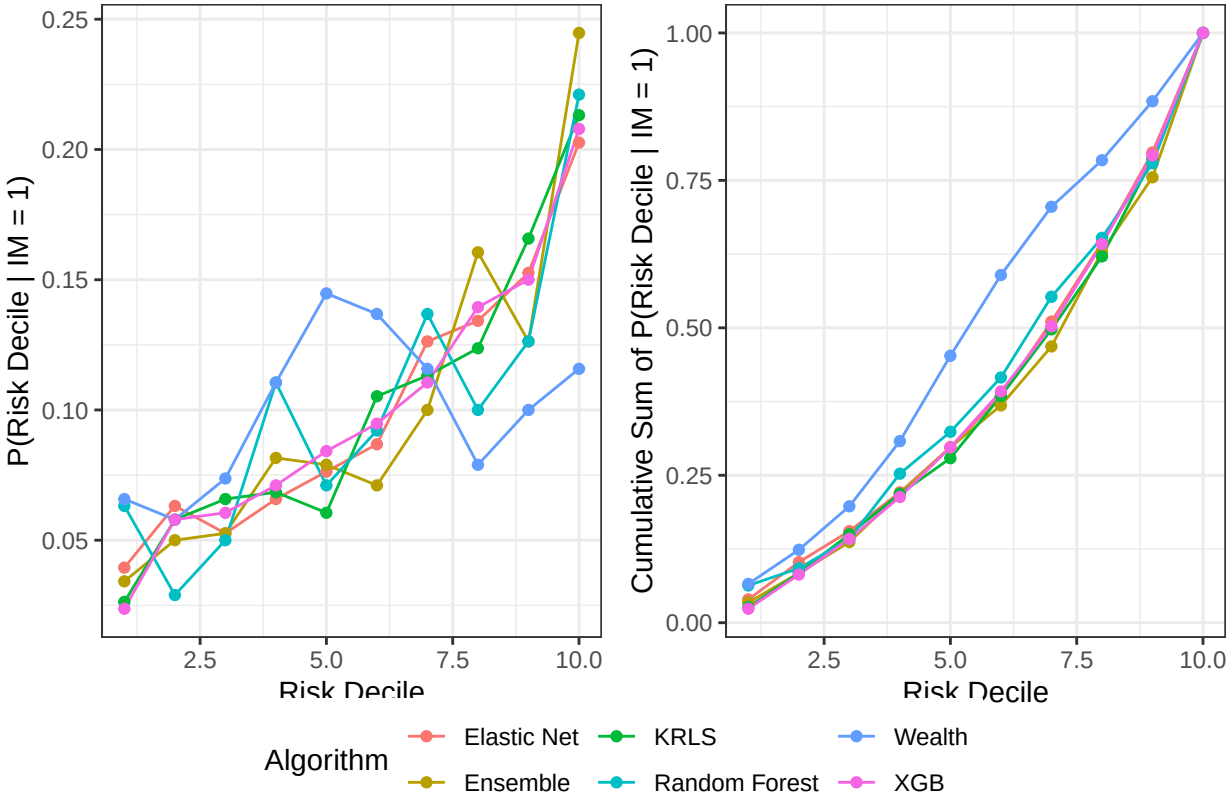

Figure 35: Probability of membership in a particular risk decile given a mortality

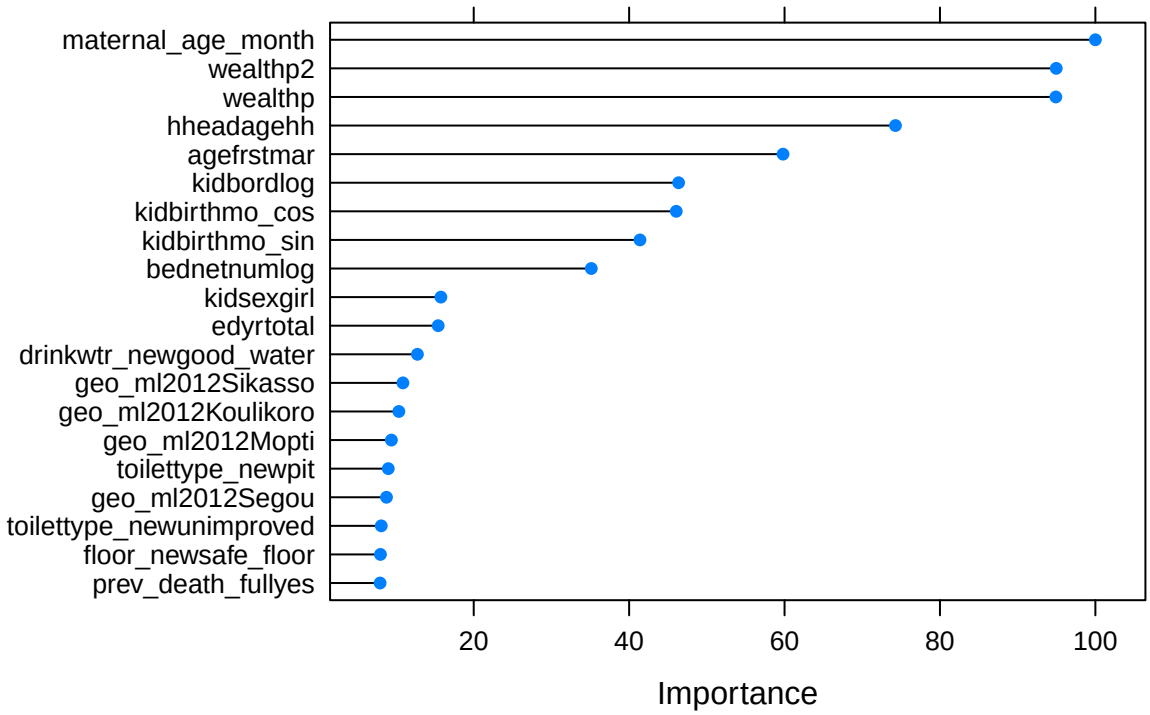

Figure 36: Variable Importance Plot generated by the Random Forest model

B.13.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.13.7 Elastic Net Coefficients

Table 104: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.3406	-2.3406
pregterminPregTerm	-0.1313	-0.0349
agefrstmar	-0.0104	-0.0414
hheadagehh	0.0012	0.0141
urbanRural	0.3267	0.1378
kidsexgirl	-0.2961	-0.148
hheadsexfemale	0.2519	0.0621
edyrtotal	-0.0201	-0.0588
drinkwtr_newgood_water	0.036	0.0171
floor_newsafe_floor	0.0548	0.0242
maternal_age_month	-8e-04	-0.0667
toilettype_newpit	-	-
toilettype_newunimproved	0.0054	0.0018
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-0.0948	-0.0237
religion_newOther	0.3268	0.0577
bednetnumlog	-0.0959	-0.0472
prev_deathyes	0.2885	0.1097
prev_death_fullyes	0.722	0.3026
kidbordlog	-0.205	-0.1388
kidbirthmo_sin	0.0559	0.0403
kidbirthmo_cos	-0.0222	-0.0151
geo_ml2012Koulikoro	-	-
geo_ml2012Sikasso	0.1817	0.0699
geo_ml2012Segou	-0.1112	-0.0435
geo_ml2012Mopti	0.0394	0.0144
geo_ml2012Bamako	-0.0041	-0.0014
wealthp	-	-
wealthp2	-0.1554	-0.0463

B.14 Mozambique

This sample was taken in the year 2011. There are 8026 observations in the dataset, and 6422 observations were used for the training set. Also, 467 observations were removed due to missing values. Finally, there were 537 deaths for the full dataset, and 430 deaths in the training set.

B.14.1 Variables Used

Table 105: Variables included in the model for Mozambique

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_mz2011	categorical
wealthp	numeric
wealthp2	numeric

B.14.2 Mortality Breakdown by Wealth Quintile

B.14.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 106: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.0147651

Table 107: Optimal Parameters for the Random Forest algorithm

mtry
10

Table 108: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	2	0.3	0	0.8	1	1

Table 109: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	70	0.0001355

B.14.4 Table of Results

Table 110: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.240	0.370	0.631	0.150	0.020	1.296	0.061	0.926	2.341
KRLS	0.207	0.323	0.591	0.111	0.010	1.148	0.062	0.924	2.023
Elastic Net	0.156	0.286	0.567	0.097	0.005	1.077	0.062	0.921	1.523
Random Forest	0.244	0.365	0.625	0.158	0.051	1.676	0.062	0.927	2.386
Wealth	0.102	0.242	0.536	0.073	0.089	2.105	0.888	0.917	1.000
XGB	0.179	0.326	0.598	0.110	0.015	1.228	0.062	0.922	1.750

Table 111: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.233	0.184	0.228	0.159	0.148
2	0.195	0.188	0.190	0.145	0.198
3	0.159	0.108	0.150	0.081	0.123
4	0.139	0.145	0.142	0.127	0.141
5	0.092	0.106	0.109	0.095	0.125
6	0.080	0.092	0.086	0.100	0.073
7	0.053	0.048	0.033	0.055	0.040
8	0.025	0.062	0.030	0.098	0.059
9	0.017	0.033	0.020	0.061	0.041
10	0.006	0.034	0.011	0.078	0.050

B.14.5 Performance Plots

Below are various plots showing the performance for each model considered in Mozambique.

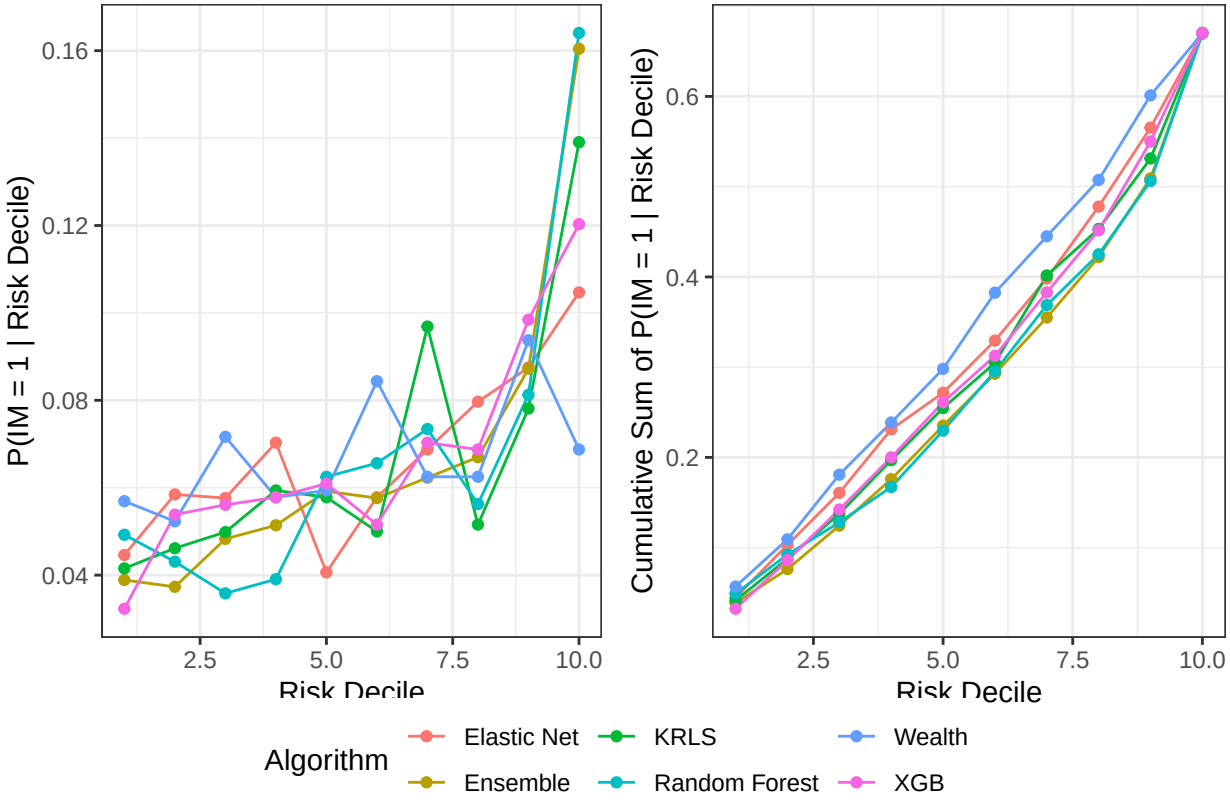

Figure 37: Probability of a mortality given that the observation is in a particular risk decile

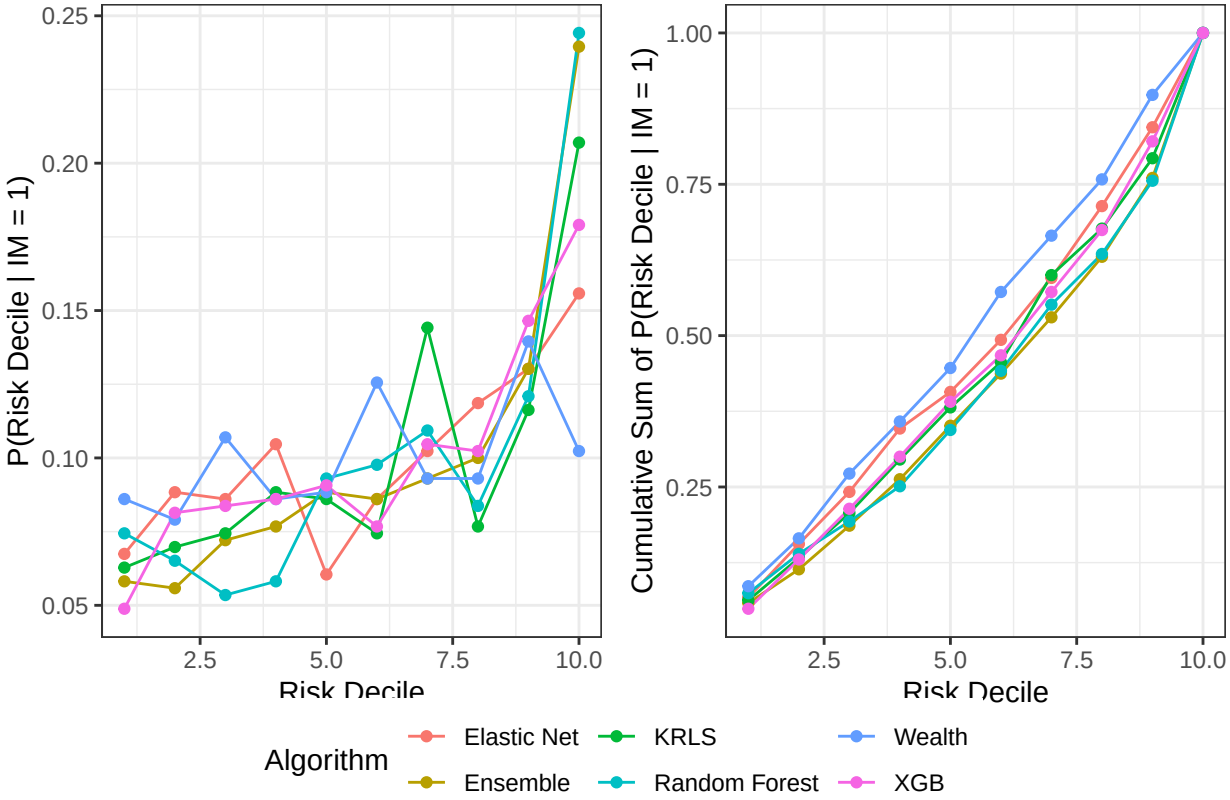

Figure 38: Probability of membership in a particular risk decile given a mortality

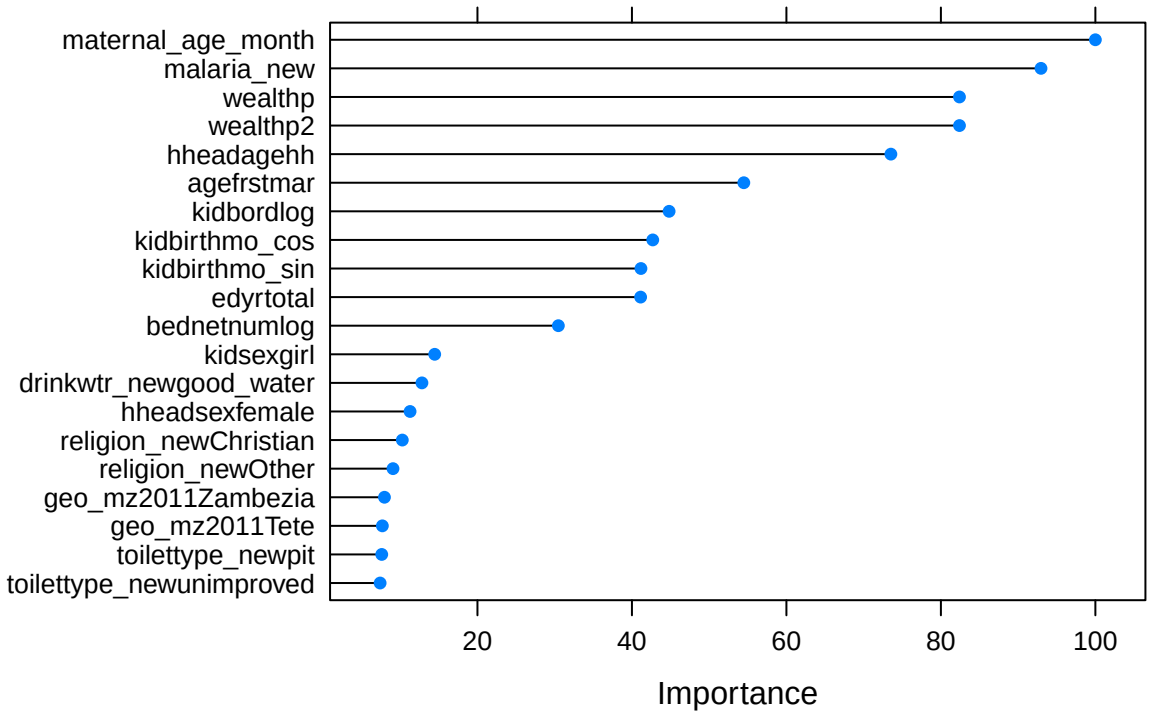

Figure 39: Variable Importance Plot generated by the Random Forest model

B.14.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.14.7 Elastic Net Coefficients

Table 112: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.5583	-2.5583
pregterminPregTerm	-	-
agefrstmar	-	-
hheadagehh	-0.0016	-0.0198
urbanRural	0.1086	0.0501
kidsexgirl	-0.0489	-0.0244
hheadsexfemale	0.0672	0.031
edyrtotal	-	-
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	-0.001	-0.085
cookfuel_newclean_fuel	-0.1191	-0.0196
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	0.0117	0.0056
bednetnumlog	-	-
prev_deathyes	0.291	0.1152
prev_death_fullyes	0.1114	0.048
malaria_new	0.3039	0.0472
kidbordlog	-0.0452	-0.0308
kidbirthmo_sin	0.1028	0.0734
kidbirthmo_cos	-	-
geo_mz2011Cabo delgado	-0.2012	-0.0578
geo_mz2011Nampula	-0.1401	-0.0389
geo_mz2011Zambezia	0.2932	0.098
geo_mz2011Tete	0.3138	0.0934
geo_mz2011Manica	-	-
geo_mz2011Sofala	-	-
geo_mz2011Inhambane	-0.1593	-0.0411
geo_mz2011Gaza	-0.0092	-0.0025
geo_mz2011Maputo Provincia	0.0655	0.0168
geo_mz2011City of Maputo	0.0165	0.0041
wealthp	-	-
wealthp2	-	-

B.15 Niger

This sample was taken in the year 2012. There are 9612 observations in the dataset, and 7690 observations were used for the training set. Also, 38 observations were removed due to missing values. Finally, there were 586 deaths for the full dataset, and 469 deaths in the training set.

B.15.1 Variables Used

Table 113: Variables included in the model for Niger

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_ne2012	categorical
wealthp	numeric
wealthp2	numeric

B.15.2 Mortality Breakdown by Wealth Quintile

B.15.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 114: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1642857	0.0106825

Table 115: Optimal Parameters for the Random Forest algorithm

mtry
8

Table 116: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 117: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	58	0.0001792

B.15.4 Table of Results

Table 118: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.194	0.335	0.635	0.150	0.018	1.287	0.056	0.925	2.021
KRLS	0.149	0.282	0.625	0.096	0.009	1.154	0.057	0.923	1.554
Elastic Net	0.149	0.307	0.612	0.086	0.008	1.127	0.057	0.923	1.554
Random Forest	0.205	0.307	0.610	0.144	0.046	1.671	0.057	0.927	2.131
Wealth	0.096	0.196	0.532	0.067	0.158	2.954	0.802	0.919	1.000
XGB	0.160	0.286	0.607	0.091	0.009	1.147	0.057	0.924	1.665

Table 119: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.134	0.116	0.135	0.121	0.105
2	0.178	0.118	0.159	0.091	0.138
3	0.134	0.152	0.148	0.139	0.146
4	0.120	0.141	0.126	0.112	0.134
5	0.139	0.114	0.134	0.085	0.133
6	0.124	0.135	0.126	0.114	0.138
7	0.085	0.090	0.099	0.090	0.105
8	0.066	0.093	0.059	0.112	0.074
9	0.016	0.030	0.011	0.079	0.022
10	0.003	0.018	0.000	0.055	0.003

B.15.5 Performance Plots

Below are various plots showing the performance for each model considered in Niger.

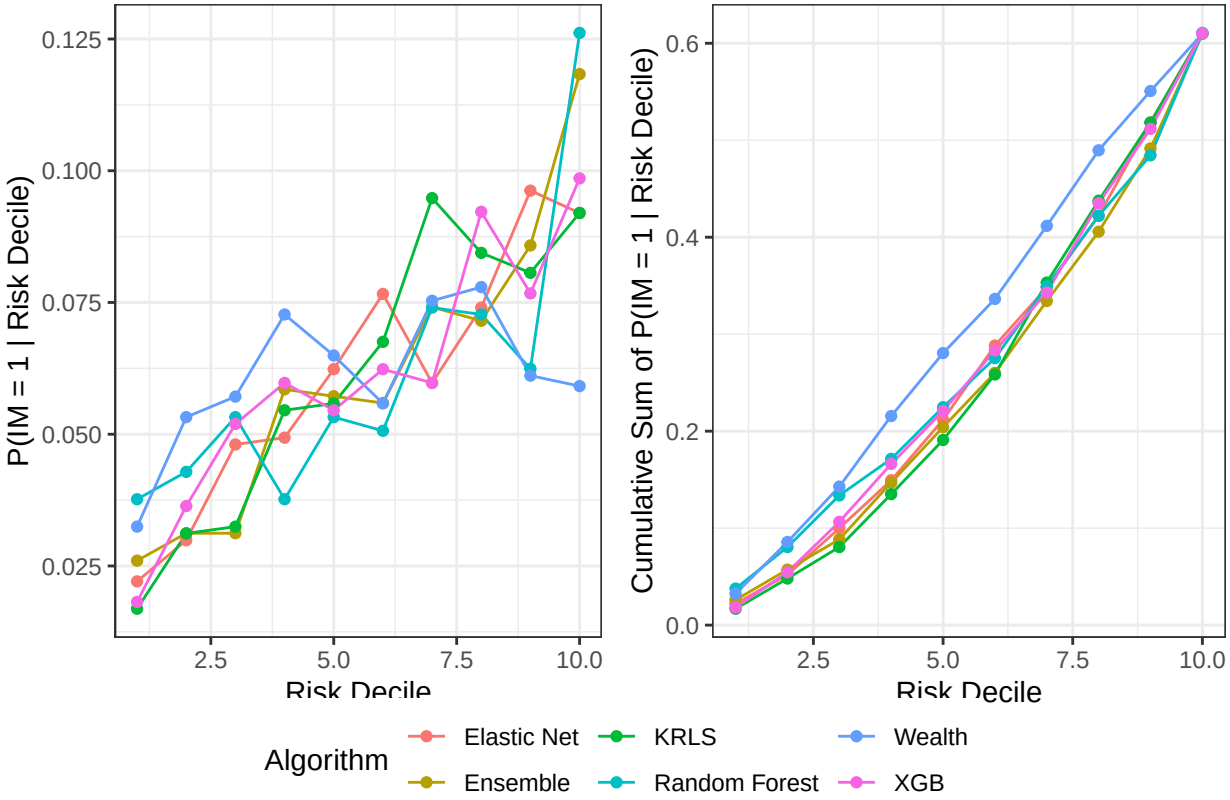

Figure 40: Probability of a mortality given that the observation is in a particular risk decile

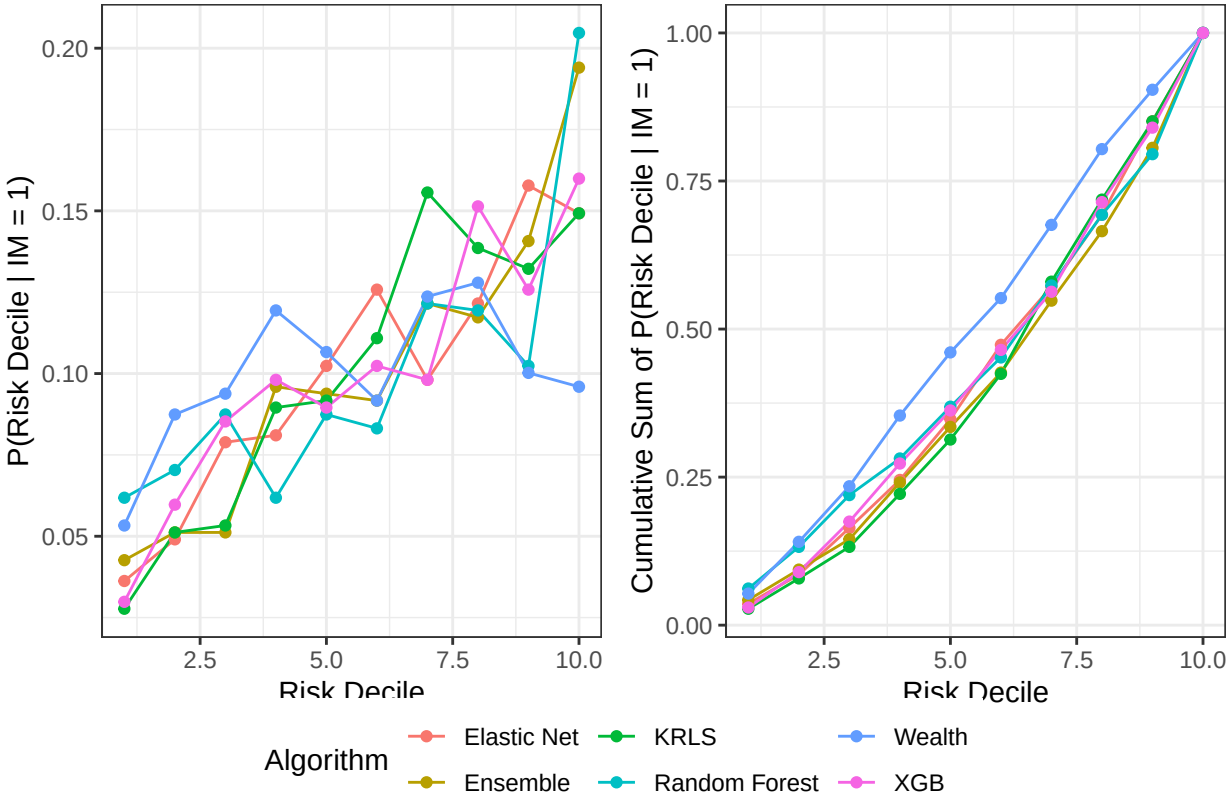

Figure 41: Probability of membership in a particular risk decile given a mortality

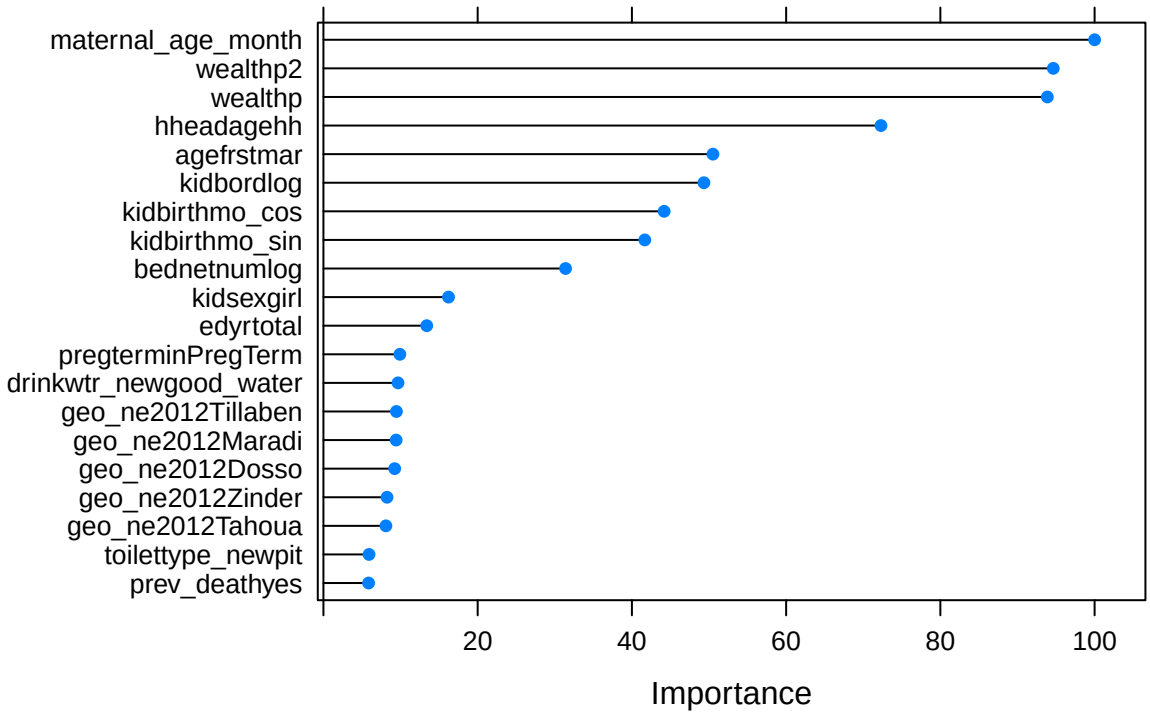

Figure 42: Variable Importance Plot generated by the Random Forest model

B.15.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.15.7 Elastic Net Coefficients

Table 120: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.654	-2.654
pregterminPregTerm	0.1776	0.0649
agefrstmar	-	-
hheadagehh	-	-
urbanRural	0.2284	0.0943
kidsexgirl	-	-
hheadsexfemale	-0.0029	-9e-04
edyrtotal	-0.0107	-0.0273
drinkwtr_newgood_water	-0.2188	-0.1032
floor_newsafe_floor	-0.0835	-0.0282
maternal_age_month	-6e-04	-0.0499
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
bednetnumlog	-	-
prev_deathyes	0.3652	0.1591
prev_death_fullyes	0.1226	0.0589
kidbordlog	-0.061	-0.0431
kidbirthmo_sin	-0.0559	-0.0402
kidbirthmo_cos	-	-
geo_ne2012Diffa	-0.644	-0.185
geo_ne2012Dosso	0.0857	0.0298
geo_ne2012Maradi	0.0059	0.0024
geo_ne2012Tahoua	-	-
geo_ne2012Tillaben	0.1678	0.0572
geo_ne2012Zinder	-	-
geo_ne2012Niamey	0.0037	0.001
wealthp	-	-
wealthp2	-	-

B.16 Nigeria

This sample was taken in the year 2013. There are 23559 observations in the dataset, and 18848 observations were used for the training set. Also, 663 observations were removed due to missing values. Finally, there were 1686 deaths for the full dataset, and 1349 deaths in the training set.

B.16.1 Variables Used

Table 121: Variables included in the model for Nigeria

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_ng2013	categorical
wealthp	numeric
wealthp2	numeric

B.16.2 Mortality Breakdown by Wealth Quintile

B.16.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 122: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1642857	0.0044743

Table 123: Optimal Parameters for the Random Forest algorithm

mtry
9

Table 124: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.4	0	0.8	1	1

Table 125: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	60	0.0001595

B.16.4 Table of Results

Table 126: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.199	0.339	0.634	0.127	0.017	1.227	0.065	0.922	1.604
KRLS	0.186	0.332	0.623	0.108	0.012	1.164	0.066	0.921	1.502
Elastic Net	0.173	0.338	0.621	0.108	0.012	1.171	0.066	0.920	1.401
Random Forest	0.175	0.294	0.603	0.117	0.031	1.375	0.067	0.920	1.413
Wealth	0.124	0.251	0.568	0.087	0.223	1.934	0.993	0.916	1.000
XGB	0.189	0.342	0.622	0.113	0.012	1.175	0.066	0.921	1.526

Table 127: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.211	0.218	0.231	0.190	0.215
2	0.221	0.193	0.240	0.137	0.222
3	0.180	0.171	0.181	0.140	0.200
4	0.155	0.170	0.163	0.153	0.168
5	0.097	0.079	0.087	0.086	0.081
6	0.049	0.050	0.046	0.071	0.038
7	0.039	0.049	0.030	0.070	0.030
8	0.029	0.036	0.015	0.059	0.031
9	0.011	0.016	0.003	0.041	0.012
10	0.006	0.019	0.003	0.052	0.003

B.16.5 Performance Plots

Below are various plots showing the performance for each model considered in Nigeria.

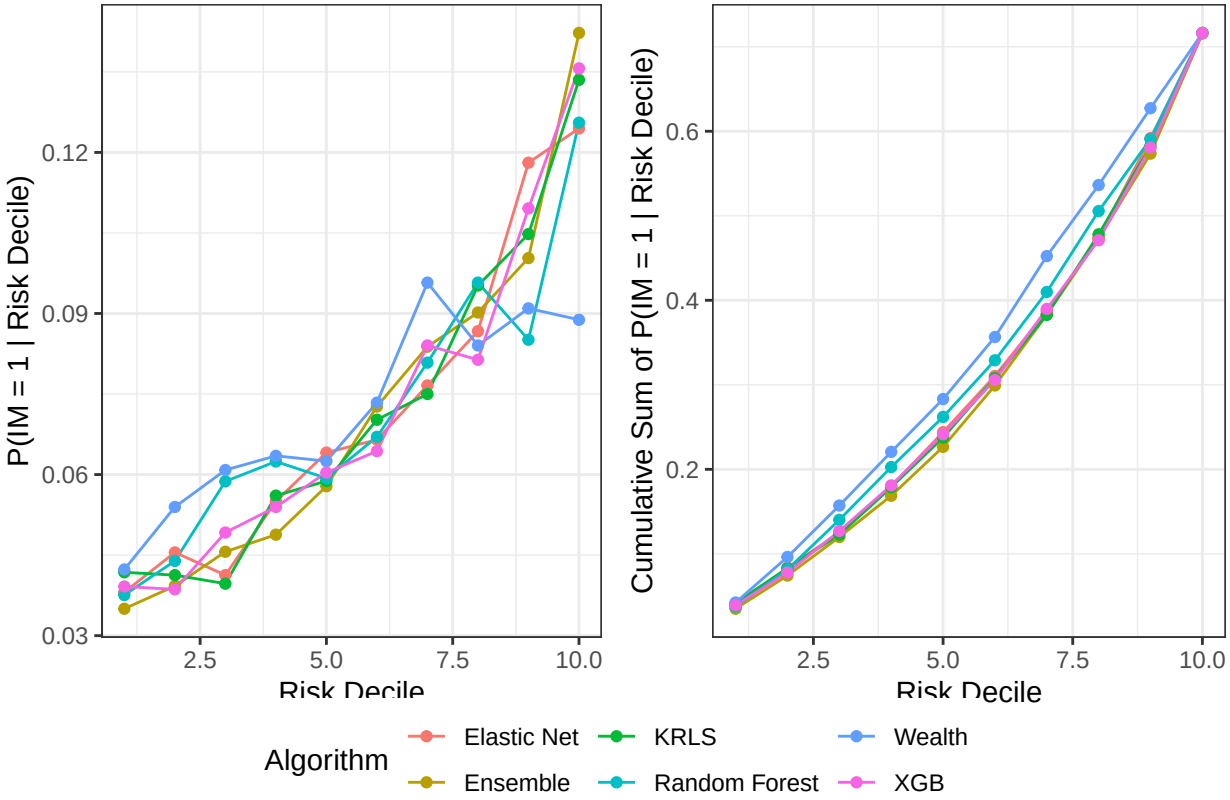

Figure 43: Probability of a mortality given that the observation is in a particular risk decile

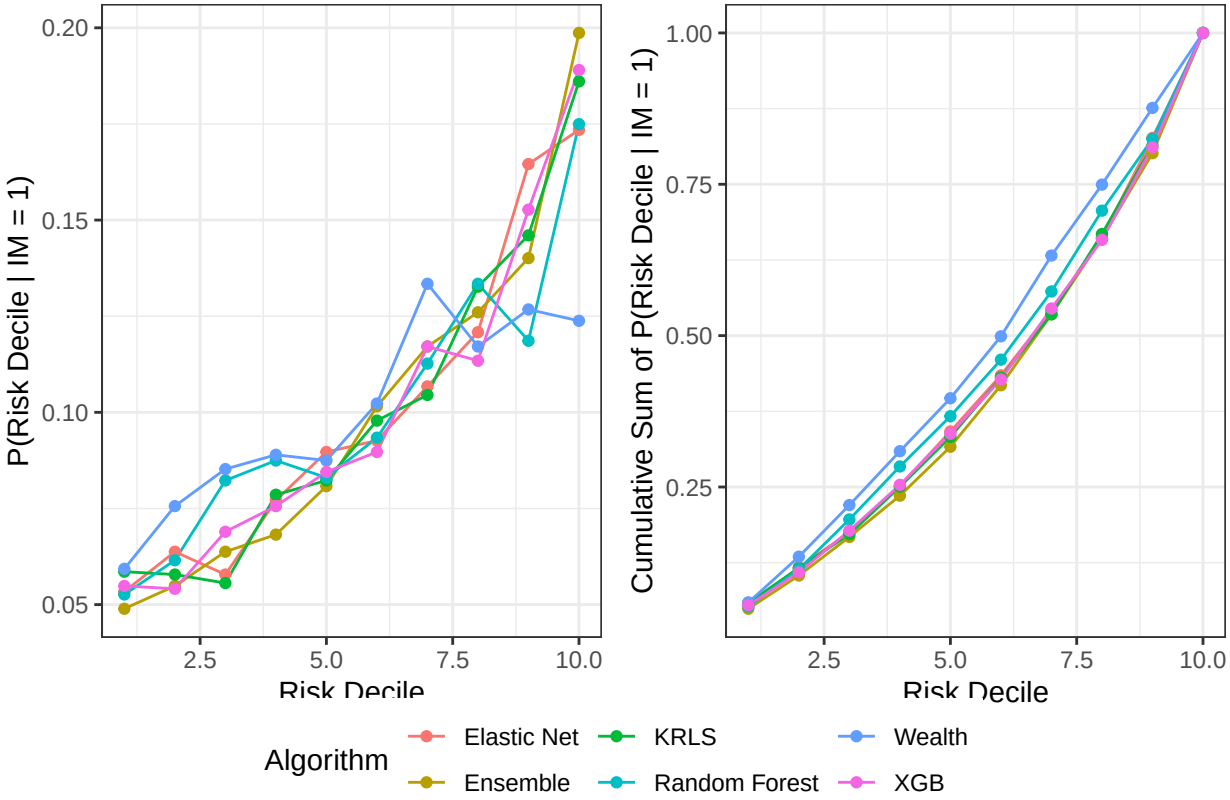

Figure 44: Probability of membership in a particular risk decile given a mortality

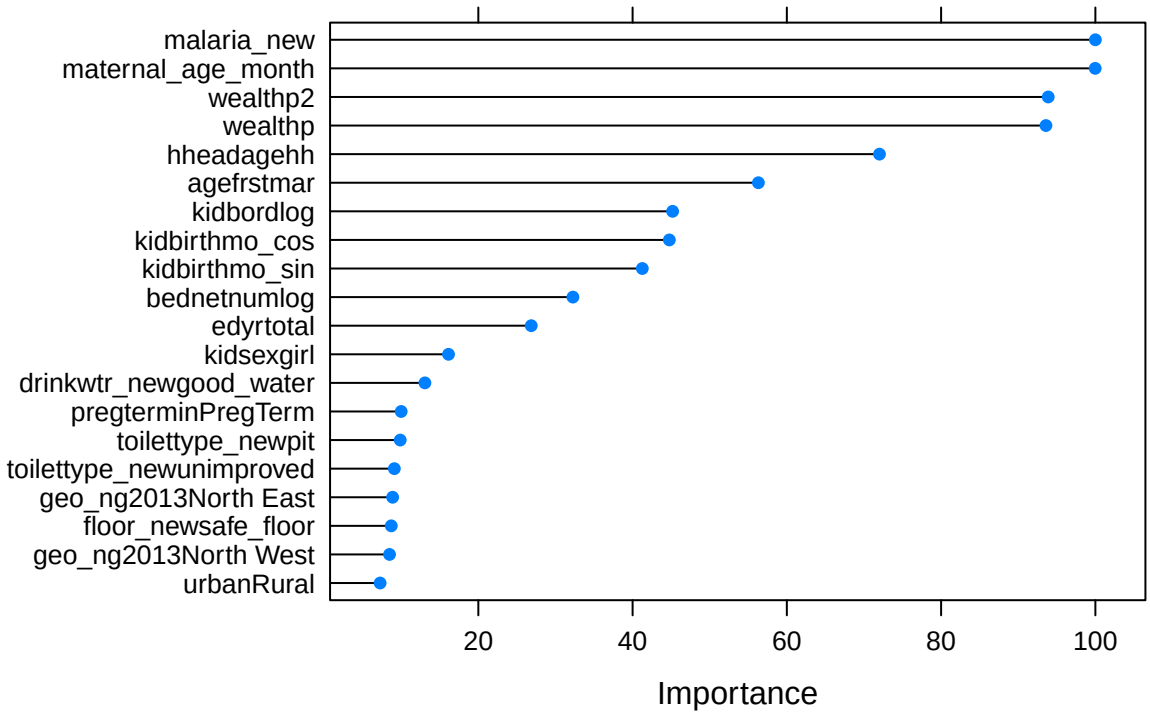

Figure 45: Variable Importance Plot generated by the Random Forest model

B.16.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.16.7 Elastic Net Coefficients

Table 128: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.6029	-2.6029
pregterminPregTerm	0.2537	0.0809
agefrstmar	0.0095	0.042
hheadagehh	-	-
urbanRural	0.1704	0.0799
kidsexgirl	-0.1129	-0.0564
hheadsexfemale	-0.0068	-0.002
edyrtotal	-0.02	-0.1048
drinkwtr_newgood_water	-0.0115	-0.0057
floor_newsafe_floor	-0.129	-0.0642
maternal_age_month	-6e-04	-0.0518
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-0.0323	-0.004
bednetnumlog	-	-
prev_deathyes	0.4483	0.1862
prev_death_fullyes	0.282	0.1308
malaria_new	0.1074	0.0116
kidbordlog	-0.1325	-0.0943
kidbirthmo_sin	-0.0957	-0.069
kidbirthmo_cos	-0.0238	-0.0164
geo_ng2013North East	-	-
geo_ng2013North West	0.1319	0.0613
geo_ng2013South East	0.3659	0.1024
geo_ng2013South South	-0.0312	-0.0099
geo_ng2013South West	-	-
wealthp	-	-
wealthp2	-0.1199	-0.0357

B.17 Rwanda

This sample was taken in the year 2014. There are 5350 observations in the dataset, and 4281 observations were used for the training set. Also, 705 observations were removed due to missing values. Finally, there were 168 deaths for the full dataset, and 135 deaths in the training set.

B.17.1 Variables Used

Table 129: Variables included in the model for Rwanda

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_rw2014	categorical
wealthp	numeric
wealthp2	numeric

B.17.2 Mortality Breakdown by Wealth Quintile

B.17.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 130: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
1	0.0048172

Table 131: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 132: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.8	1	1

Table 133: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	56	0.0003284

B.17.4 Table of Results

Table 134: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.186	0.320	0.618	0.068	0.005	1.172	0.030	0.936	1.251
KRLS	0.142	0.312	0.609	0.052	0.003	1.099	0.030	0.935	0.952
Elastic Net	0.164	0.357	0.623	0.048	0.004	1.114	0.030	0.936	1.103
Random Forest	0.207	0.296	0.578	0.068	0.010	1.517	0.031	0.938	1.387
Wealth	0.149	0.259	0.561	0.041	0.196	5.841	0.814	0.936	1.000
XGB	0.126	0.266	0.556	0.042	0.003	1.090	0.031	0.935	0.845

Table 135: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.255	0.293	0.252	0.267	0.255
2	0.182	0.175	0.213	0.117	0.220
3	0.171	0.098	0.164	0.062	0.076
4	0.117	0.071	0.088	0.064	0.067
5	0.074	0.079	0.095	0.081	0.071
6	0.079	0.098	0.083	0.086	0.124
7	0.035	0.096	0.049	0.107	0.091
8	0.033	0.067	0.033	0.088	0.076
9	0.029	0.017	0.012	0.064	0.000
10	0.021	0.021	0.005	0.060	0.014

B.17.5 Performance Plots

Below are various plots showing the performance for each model considered in Rwanda.

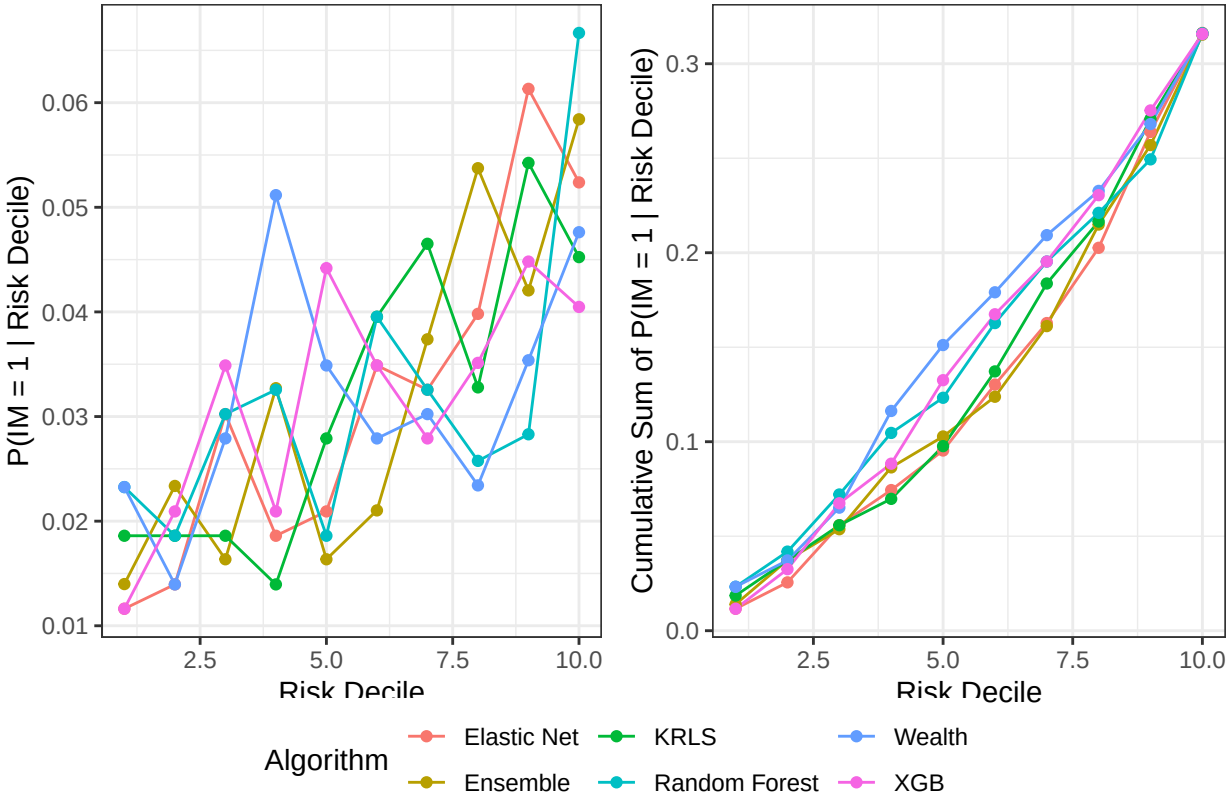

Figure 46: Probability of a mortality given that the observation is in a particular risk decile

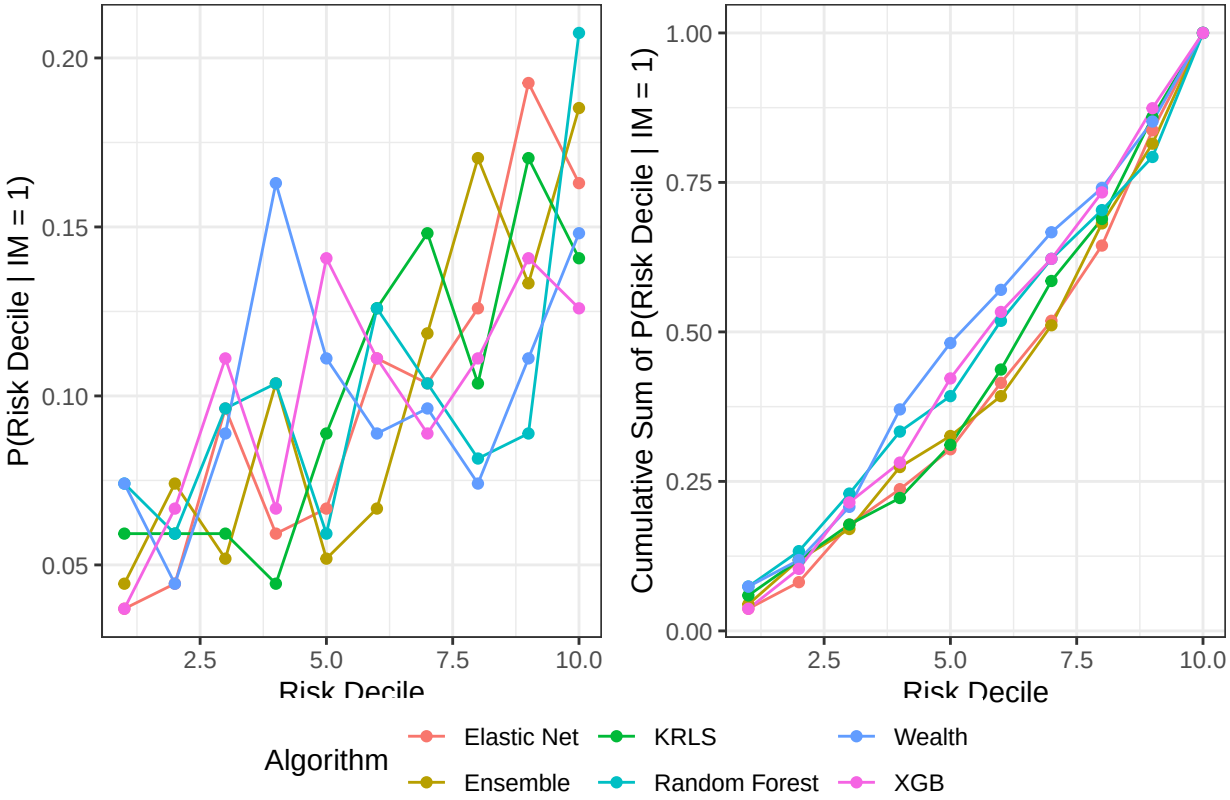

Figure 47: Probability of membership in a particular risk decile given a mortality

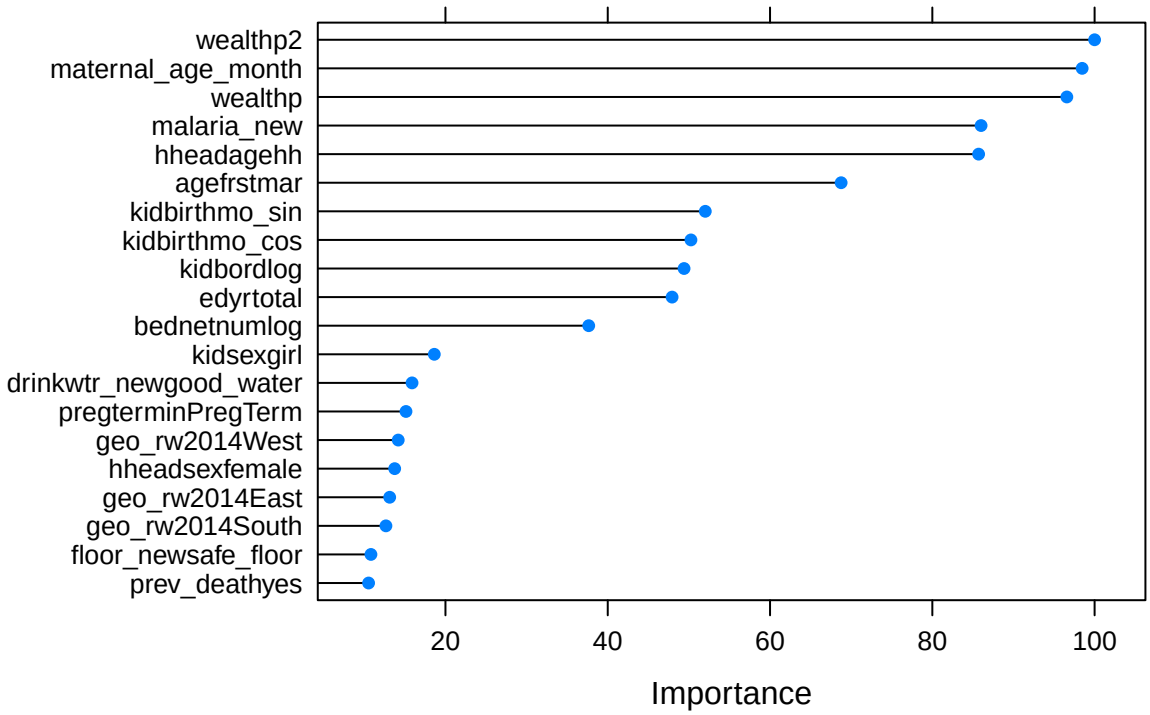

Figure 48: Variable Importance Plot generated by the Random Forest model

B.17.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.17.7 Elastic Net Coefficients

Table 136: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.9572	-2.9572
pregterminPregTerm	0.0025	9e-04
agefrstmar	-	-
hheadagehh	-	-
urbanRural	-	-
kidsexgirl	-	-
hheadsexfemale	-	-
edyrtotal	-0.0044	-0.0158
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-0.5674	-0.2638
prev_deathyes	-	-
prev_death_fullyes	0.148	0.0587
malaria_new	-	-
kidbordlog	-	-
kidbirthmo_sin	-	-
kidbirthmo_cos	-	-
geo_rw2014South	-	-
geo_rw2014West	-	-
geo_rw2014North	-0.0035	-0.0012
geo_rw2014East	-	-
wealthp	-	-
wealthp2	-	-

B.18 Senegal

This sample was taken in the year 2017. There are 9371 observations in the dataset, and 7497 observations were used for the training set. Also, 214 observations were removed due to missing values. Finally, there were 401 deaths for the full dataset, and 321 deaths in the training set.

B.18.1 Variables Used

Table 137: Variables included in the model for Senegal

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_sn2017	categorical
wealthp	numeric
wealthp2	numeric

B.18.2 Mortality Breakdown by Wealth Quintile

B.18.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 138: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.0127515

Table 139: Optimal Parameters for the Random Forest algorithm

mtry
11

Table 140: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
150	5	0.3	0	0.6	1	1

Table 141: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	74	0.0002243

B.18.4 Table of Results

Table 142: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.224	0.339	0.609	0.148	0.033	1.810	0.039	0.933	2.660
KRLS	0.165	0.315	0.581	0.075	0.004	1.100	0.041	0.931	1.957
Elastic Net	0.149	0.299	0.580	0.071	0.004	1.101	0.041	0.930	1.773
Random Forest	0.209	0.311	0.587	0.144	0.059	2.194	0.040	0.933	2.475
Wealth	0.084	0.203	0.535	0.045	0.106	1.687	0.956	0.927	1.000
XGB	0.209	0.333	0.595	0.121	0.064	3.297	0.041	0.933	2.476

Table 143: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.181	0.100	0.146	0.106	0.067
2	0.147	0.116	0.159	0.116	0.105
3	0.126	0.146	0.135	0.130	0.134
4	0.122	0.114	0.127	0.096	0.115
5	0.103	0.103	0.115	0.090	0.111
6	0.110	0.088	0.120	0.091	0.076
7	0.085	0.112	0.096	0.108	0.135
8	0.050	0.100	0.052	0.088	0.127
9	0.052	0.067	0.043	0.075	0.075
10	0.024	0.055	0.005	0.099	0.054

B.18.5 Performance Plots

Below are various plots showing the performance for each model considered in Senegal.

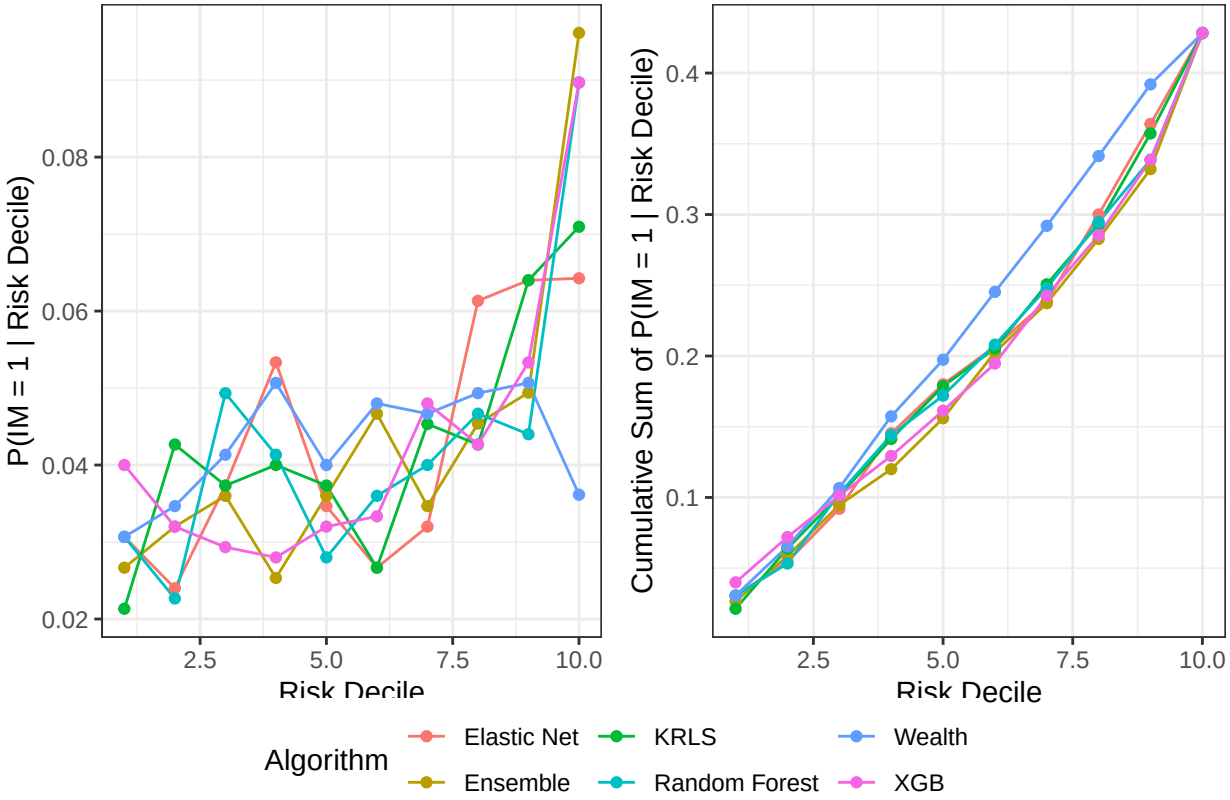

Figure 49: Probability of a mortality given that the observation is in a particular risk decile

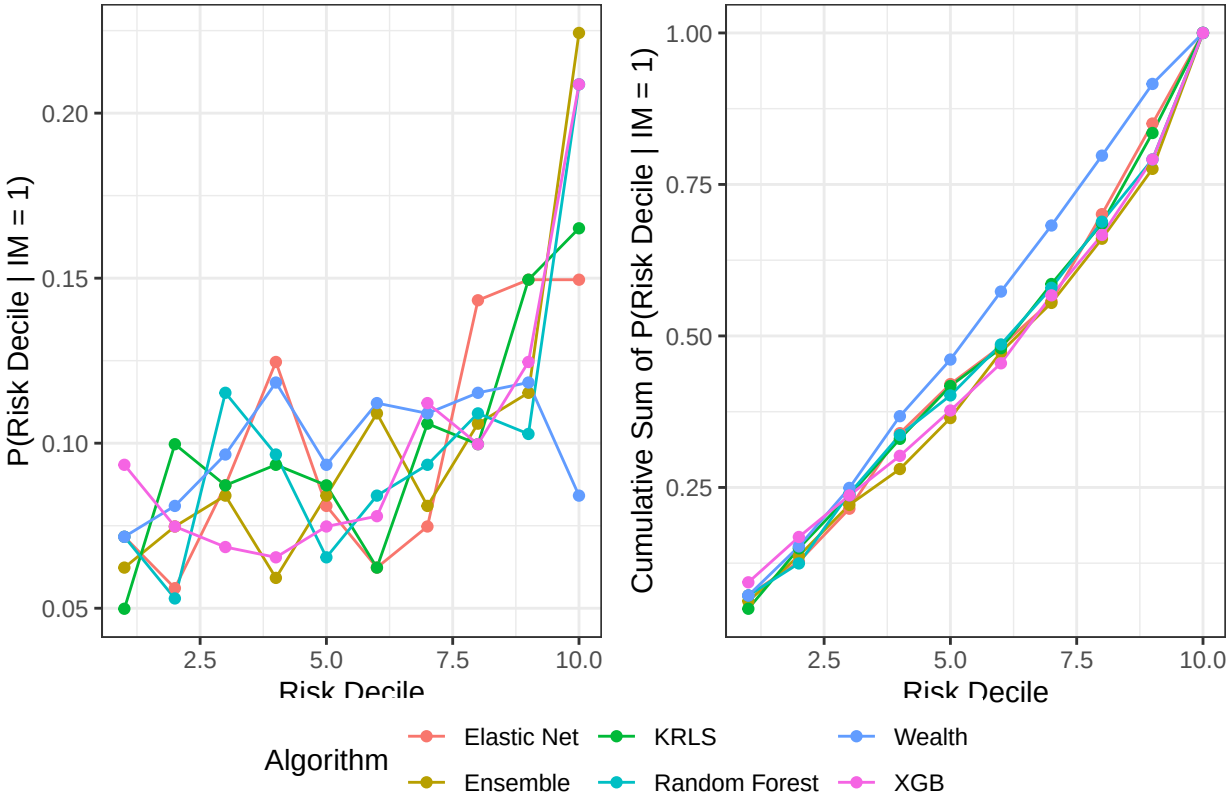

Figure 50: Probability of membership in a particular risk decile given a mortality

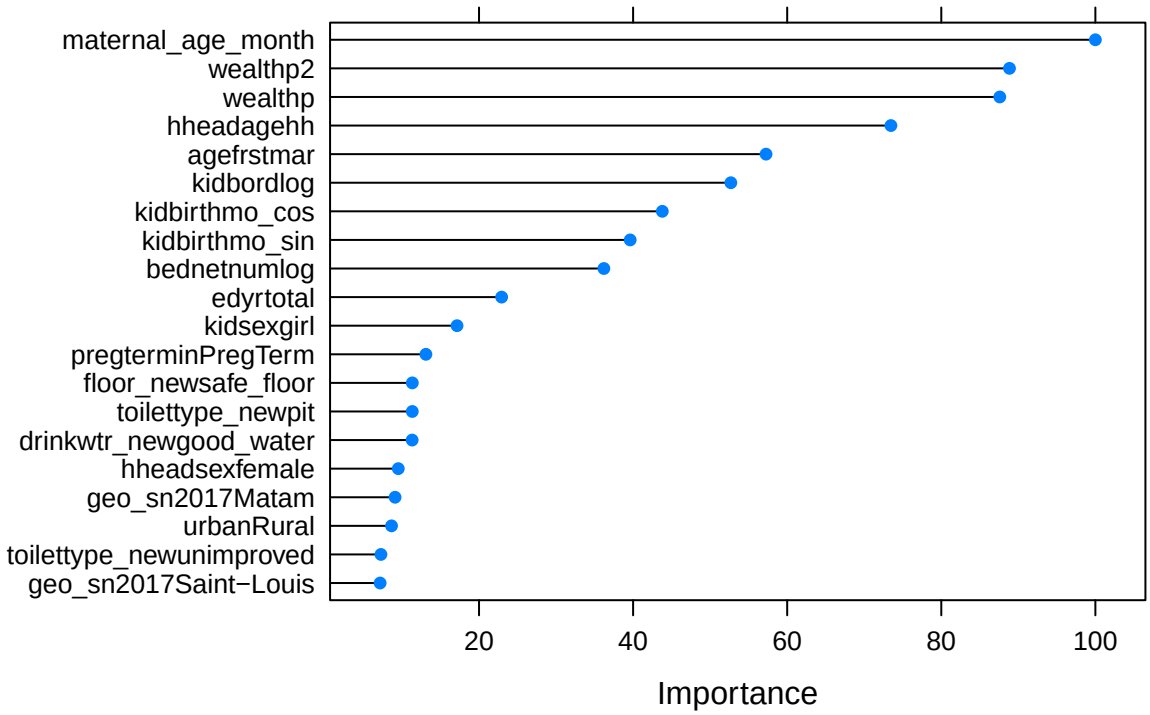

Figure 51: Variable Importance Plot generated by the Random Forest model

B.18.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.18.7 Elastic Net Coefficients

Table 144: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.8655	-2.8655
pregterminPregTerm	0.1831	0.0729
agefrstmar	-	-
hheadagehh	-2e-04	-0.0031
urbanRural	0.0245	0.0116
kidsexgirl	-0.0474	-0.0237
hheadsexfemale	-0.0105	-0.0044
edyrtotal	-0.0123	-0.0432
drinkwtr_newgood_water	-0.0767	-0.0361
floor_newsafe_floor	-0.1394	-0.0695
maternal_age_month	-5e-04	-0.0419
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-0.2178	-0.0313
religion_newHindu	-	-
religion_newMuslim	0.0229	0.0033
religion_newOther	1.2575	0.0325
bednetnumlog	-0.0585	-0.0358
prev_deathyes	0.3571	0.1207
prev_death_fullyes	0.1783	0.0703
kidbordlog	-	-
kidbirthmo_sin	0.0883	0.0623
kidbirthmo_cos	-0.1382	-0.0974
geo_sn2017Ziguinchor	-0.0471	-0.0095
geo_sn2017Diourbel	-	-
geo_sn2017Saint-Louis	0.1278	0.0311
geo_sn2017Tambacounda	-0.0451	-0.0124
geo_sn2017Kaolack	-0.1627	-0.0395
geo_sn2017Thiès	0.0028	7e-04
geo_sn2017Louga	0.2438	0.063
geo_sn2017Fatick	-0.0746	-0.0196
geo_sn2017Kolda	-0.076	-0.02
geo_sn2017Matam	0.4622	0.1264
geo_sn2017Kaffrine	-	-
geo_sn2017Kedougou	-	-
geo_sn2017Sedhiou	-	-
wealthp	-	-
wealthp2	-	-

B.19 Tanzania

This sample was taken in the year 2015. There are 7688 observations in the dataset, and 6152 observations were used for the training set. Also, 319 observations were removed due to missing values. Finally, there were 319 deaths for the full dataset, and 256 deaths in the training set.

B.19.1 Variables Used

Table 145: Variables included in the model for Tanzania

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
toilettype_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
district	categorical
wealthp	numeric
wealthp2	numeric

B.19.2 Mortality Breakdown by Wealth Quintile

B.19.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 146: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.8071429	0.0030448

Table 147: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 148: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	0.75

Table 149: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	60	0.0001249

B.19.4 Table of Results

Table 150: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.188	0.262	0.582	0.077	0.005	1.145	0.04	0.932	2.169
KRLS	0.144	0.254	0.549	0.076	0.004	1.098	0.04	0.931	1.667
Elastic Net	0.121	0.226	0.561	0.066	0.002	1.054	0.04	0.930	1.397
Random Forest	0.180	0.309	0.584	0.072	0.011	1.459	0.04	0.932	2.080
Wealth	0.086	0.161	0.545	0.039	-0.128	0.251	0.95	0.928	1.000
XGB	0.164	0.285	0.560	0.065	0.005	1.110	0.04	0.932	1.895

Table 151: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.044	0.052	0.026	0.107	0.054
2	0.050	0.059	0.031	0.065	0.067
3	0.048	0.061	0.048	0.057	0.084
4	0.056	0.057	0.041	0.069	0.092
5	0.075	0.062	0.062	0.061	0.077
6	0.074	0.052	0.074	0.051	0.054
7	0.099	0.107	0.098	0.098	0.081
8	0.120	0.143	0.149	0.120	0.138
9	0.184	0.146	0.234	0.118	0.115
10	0.249	0.267	0.236	0.254	0.238

B.19.5 Performance Plots

Below are various plots showing the performance for each model considered in Tanzania.

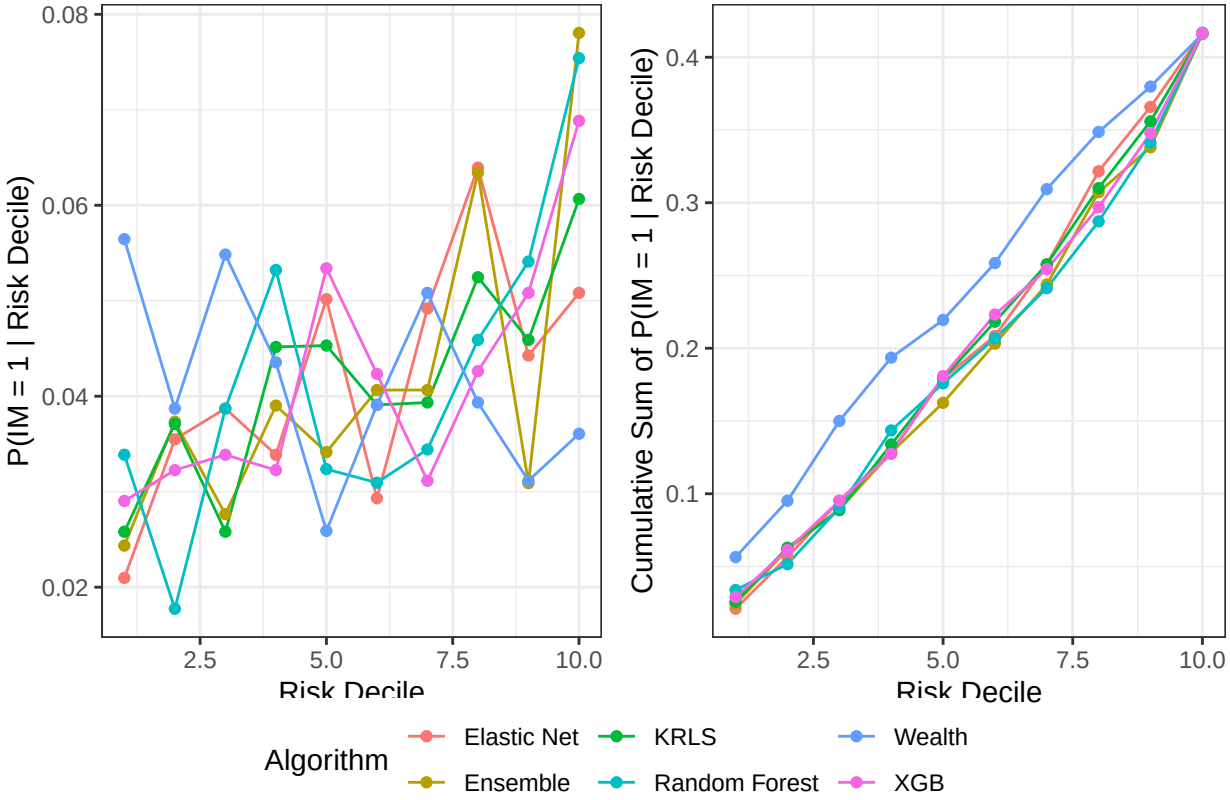

Figure 52: Probability of a mortality given that the observation is in a particular risk decile

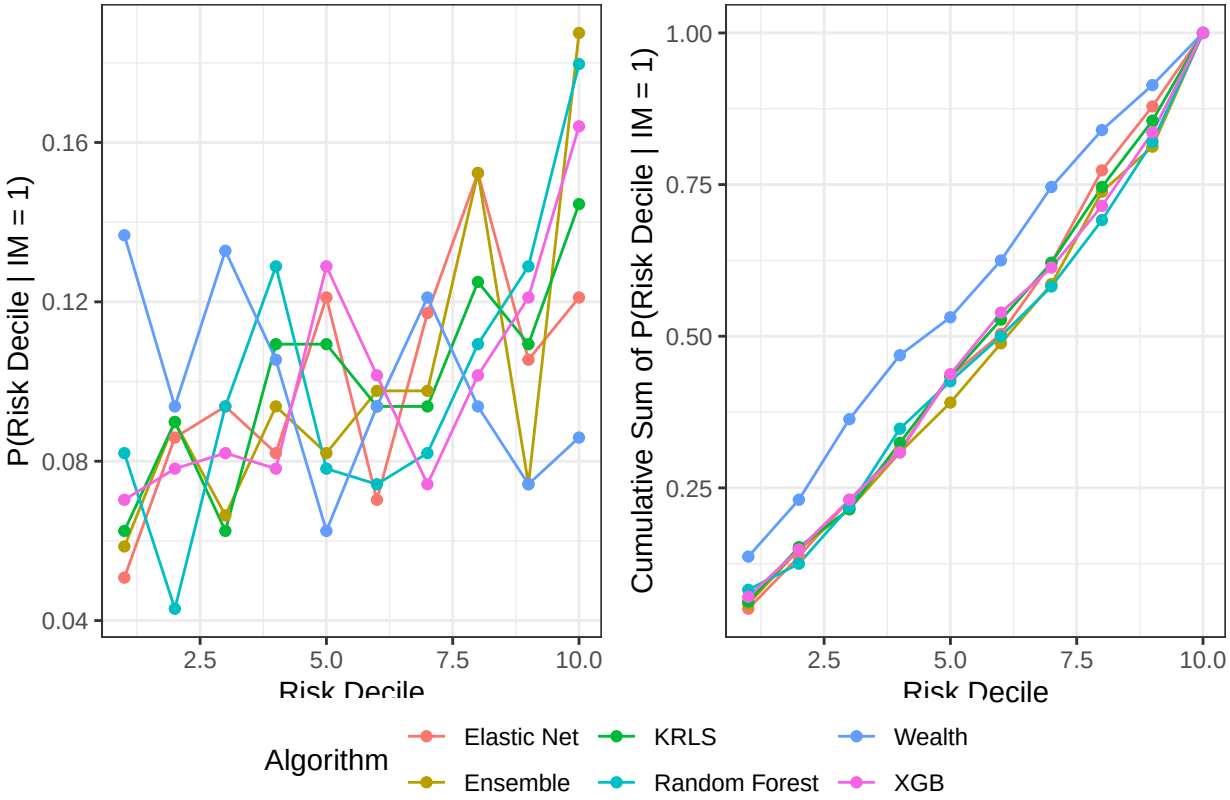

Figure 53: Probability of membership in a particular risk decile given a mortality

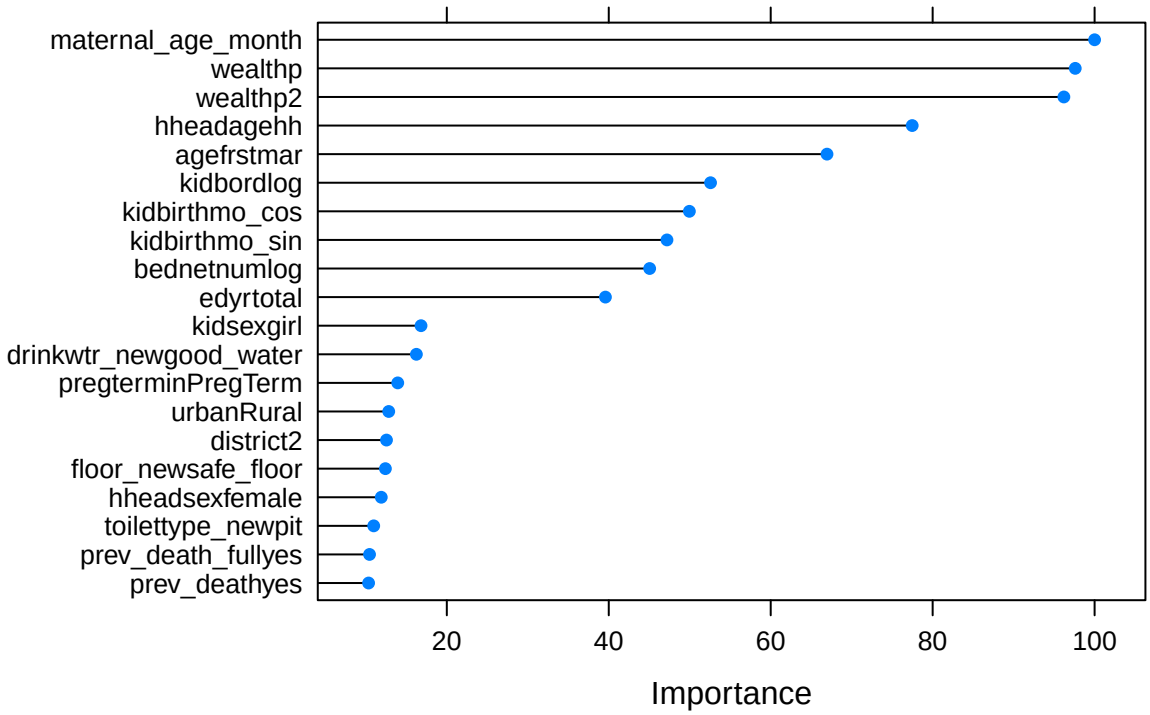

Figure 54: Variable Importance Plot generated by the Random Forest model

B.19.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.19.7 Elastic Net Coefficients

Table 152: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.8753	-2.8753
pregterminPregTerm	-	-
agefrstmar	-	-
hheadagehh	-	-
urbanRural	-0.1189	-0.05
kidsexgirl	-0.2468	-0.1234
hheadsexfemale	-	-
edyrtotal	0.0081	0.0294
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-0.4548	-0.1633
bednetnumlog	-0.0269	-0.0173
prev_deathyes	0.2938	0.1087
prev_death_fullyes	-	-
kidbordlog	-0.0745	-0.0533
kidbirthmo_sin	-	-
kidbirthmo_cos	-0.1085	-0.0768
district2	-	-
district3	-0.0047	-0.0017
district4	-0.0435	-0.0145
district5	-	-
district6	-	-
district7	0.0894	0.0178
district8	-	-
district9	0.1317	0.0099
district10	-	-
wealthp	-	-
wealthp2	-	-

B.20 Uganda

This sample was taken in the year 2016. There are 11619 observations in the dataset, and 9296 observations were used for the training set. Also, 470 observations were removed due to missing values. Finally, there were 484 deaths for the full dataset, and 388 deaths in the training set.

B.20.1 Variables Used

Table 153: Variables included in the model for Uganda

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
district	categorical
wealthp	numeric
wealthp2	numeric

B.20.2 Mortality Breakdown by Wealth Quintile

B.20.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 154: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.8071429	0.0062481

Table 155: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 156: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 157: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	268	0.0002704

B.20.4 Table of Results

Table 158: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.178	0.307	0.579	0.069	0.002	1.059	0.040	0.932	1.725
KRLS	0.134	0.237	0.535	0.062	0.001	1.027	0.040	0.930	1.299
Elastic Net	0.145	0.240	0.536	0.062	0.001	1.019	0.040	0.930	1.402
Random Forest	0.167	0.278	0.554	0.061	0.001	1.620	0.041	0.931	1.624
Wealth	0.103	0.224	0.537	0.049	-0.021	0.919	0.900	0.929	1.000
XGB	0.157	0.294	0.576	0.065	0.004	1.104	0.040	0.931	1.526

Table 159: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.229	0.165	0.235	0.150	0.125
2	0.136	0.094	0.075	0.091	0.099
3	0.117	0.081	0.082	0.107	0.086
4	0.112	0.092	0.076	0.081	0.097
5	0.082	0.087	0.081	0.091	0.095
6	0.097	0.083	0.104	0.082	0.071
7	0.086	0.090	0.105	0.093	0.086
8	0.068	0.129	0.103	0.116	0.143
9	0.046	0.103	0.087	0.109	0.102
10	0.026	0.079	0.051	0.080	0.095

B.20.5 Performance Plots

Below are various plots showing the performance for each model considered in Uganda.

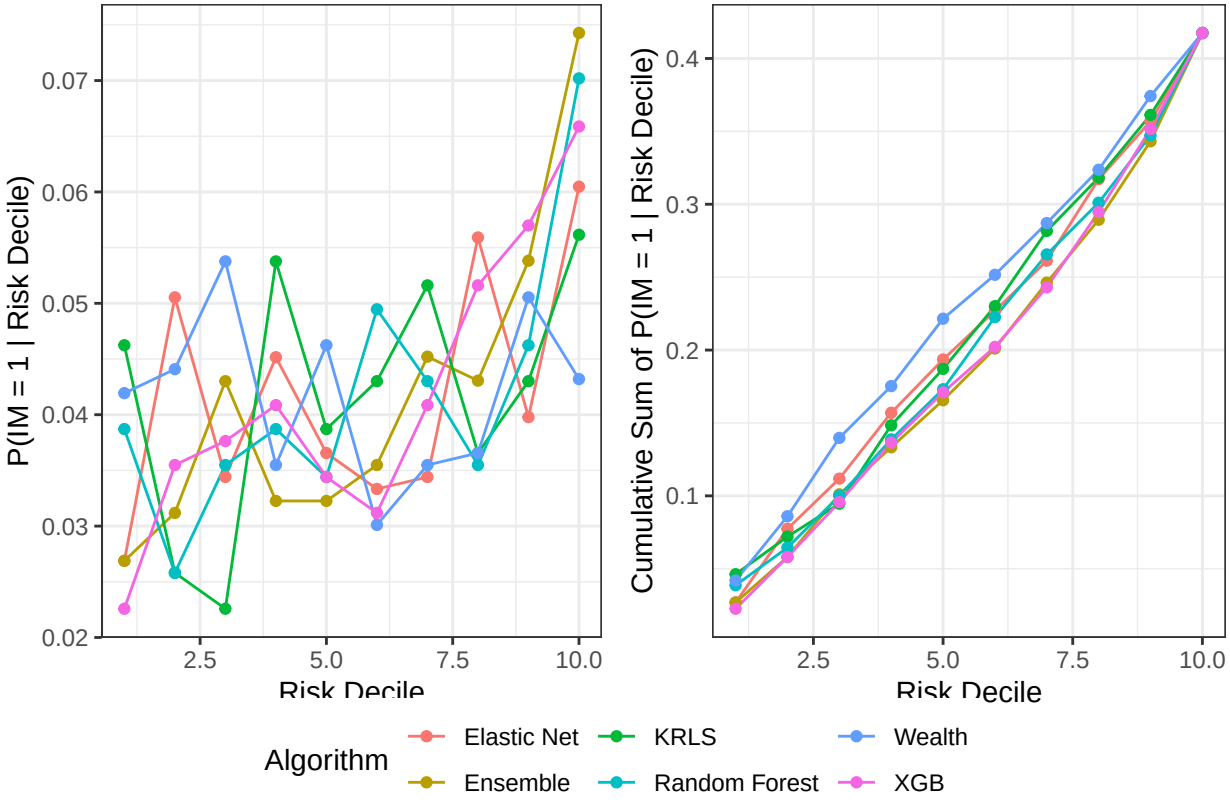

Figure 55: Probability of a mortality given that the observation is in a particular risk decile

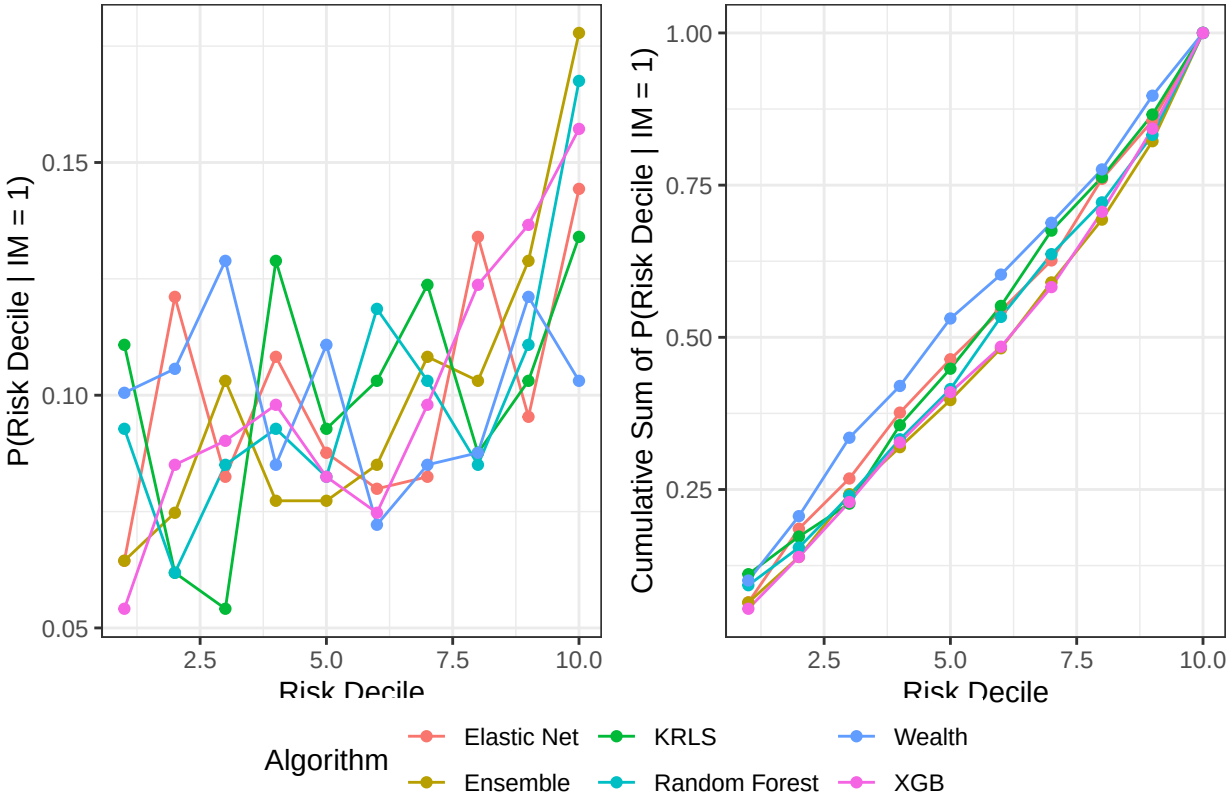

Figure 56: Probability of membership in a particular risk decile given a mortality

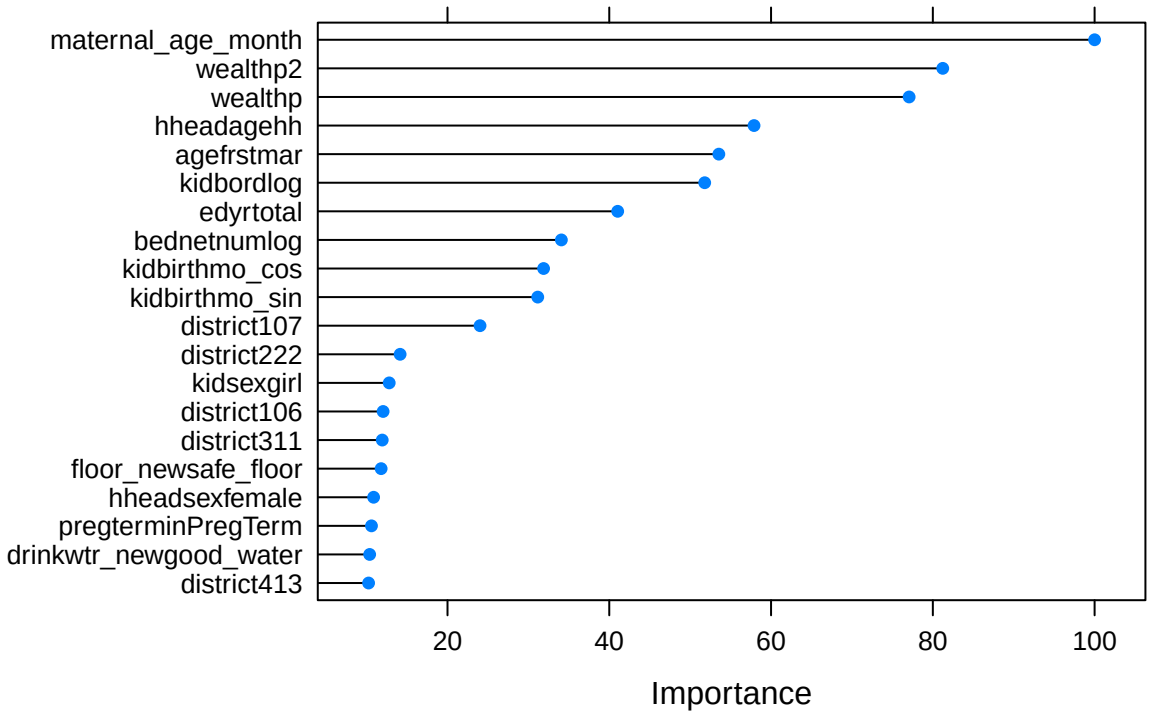

Figure 57: Variable Importance Plot generated by the Random Forest model

B.20.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.20.7 Elastic Net Coefficients

Table 160: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-3.0915	-3.0915
pregterminPregTerm	-	-
agefrstmar	-	-
hheadagehh	-	-
urbanRural	-	-
kidsexgirl	-0.0042	-0.0021
hheadsexfemale	-	-
edyrtotal	-0.0086	-0.0331
drinkwtr_newgood_water	-	-
floor_newsafe_floor	-	-
maternal_age_month	-	-
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-	-
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-	-
prev_deathyes	-	-
prev_death_fullyes	-	-
kidbordlog	-	-
kidbirthmo_sin	-	-
kidbirthmo_cos	-	-
district102	-	-
district103	-	-
district104	-	-
district105	-	-
district106	-	-
district107	0.7873	0.0828
district108	-	-
district109	-	-
district110	-	-
district111	-	-
district112	-	-
district113	-	-
district114	-	-
district115	-	-
district116	-	-
district117	-	-

Table 160: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation. *(continued)*

	Coefficients	**Standardized Coefficients**
district118	-	-
district119	-	-
district120	-	-
district121	-	-
district122	-	-
district123	-	-
district124	-	-
district201	-	-
district202	-	-
district203	-	-
district204	-	-
district205	-	-
district206	-	-
district207	-	-
district208	-	-
district209	-	-
district210	-	-
district211	-	-
district212	-	-
district213	-	-
district214	-	-
district215	-	-
district216	-	-
district217	-	-
district218	-	-
district219	-	-
district220	-	-
district221	-	-
district222	0.4808	0.0362
district223	-	-
district224	-	-
district225	-	-
district226	-	-
district227	-	-
district228	-	-
district229	-	-
district230	-	-
district231	-	-
district232	-	-
district301	-	-
district302	-	-

Table 160: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation. *(continued)*

	Coefficients	**Standardized Coefficients**
district303	-	-
district304	-	-
district305	-	-
district306	-	-
district307	-	-
district308	-	-
district309	-	-
district310	-	-
district311	-	-
district312	-	-
district313	-	-
district314	-	-
district315	-	-
district316	-	-
district317	-	-
district318	-	-
district319	-	-
district320	-	-
district321	-	-
district322	-	-
district323	-	-
district324	-	-
district325	-	-
district326	-	-
district327	-	-
district328	-	-
district329	-	-
district330	-	-
district401	-	-
district402	-	-
district403	-	-
district404	-	-
district405	-	-
district406	-	-
district407	-	-
district408	-	-
district409	-	-
district410	-	-
district411	-	-
district412	-	-
district413	0.0056	5e-04

Table 160: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation. *(continued)*

	Coefficients	**Standardized Coefficients**
district414	-	-
district415	-	-
district416	-	-
district417	-	-
district418	-	-
district419	-	-
district420	-	-
district421	-	-
district422	-	-
district423	-	-
district424	-	-
district425	-	-
district426	-	-
wealthp	-	-
wealthp2	-	-

B.21 Zambia

This sample was taken in the year 2013. There are 9811 observations in the dataset, and 7850 observations were used for the training set. Also, 835 observations were removed due to missing values. Finally, there were 483 deaths for the full dataset, and 387 deaths in the training set.

B.21.1 Variables Used

Table 161: Variables included in the model for Zambia

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
malaria_new	numeric
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
wealthp	numeric
wealthp2	numeric

B.21.2 Mortality Breakdown by Wealth Quintile

B.21.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 162: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
0.1	0.004132

Table 163: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 164: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
50	1	0.3	0	0.6	1	1

Table 165: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	50	0.0001824

B.21.4 Table of Results

Table 166: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.209	0.354	0.617	0.094	0.008	1.190	0.046	0.930	2.619
KRLS	0.186	0.333	0.588	0.083	0.005	1.098	0.047	0.930	2.327
Elastic Net	0.150	0.310	0.579	0.070	0.005	1.099	0.047	0.928	1.875
Random Forest	0.209	0.300	0.598	0.101	0.016	1.588	0.047	0.931	2.618
Wealth	0.080	0.163	0.529	0.051	-0.009	0.990	0.817	0.924	1.000
XGB	0.194	0.320	0.595	0.080	0.008	1.160	0.047	0.930	2.422

Table 167: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.071	0.117	0.095	0.149	0.100
2	0.079	0.092	0.103	0.082	0.090
3	0.072	0.087	0.083	0.074	0.113
4	0.078	0.078	0.087	0.082	0.088
5	0.063	0.094	0.072	0.106	0.112
6	0.104	0.118	0.095	0.108	0.128
7	0.130	0.101	0.106	0.074	0.102
8	0.142	0.105	0.123	0.073	0.112
9	0.162	0.128	0.141	0.128	0.101
10	0.099	0.086	0.094	0.123	0.053

B.21.5 Performance Plots

Below are various plots showing the performance for each model considered in Zambia.

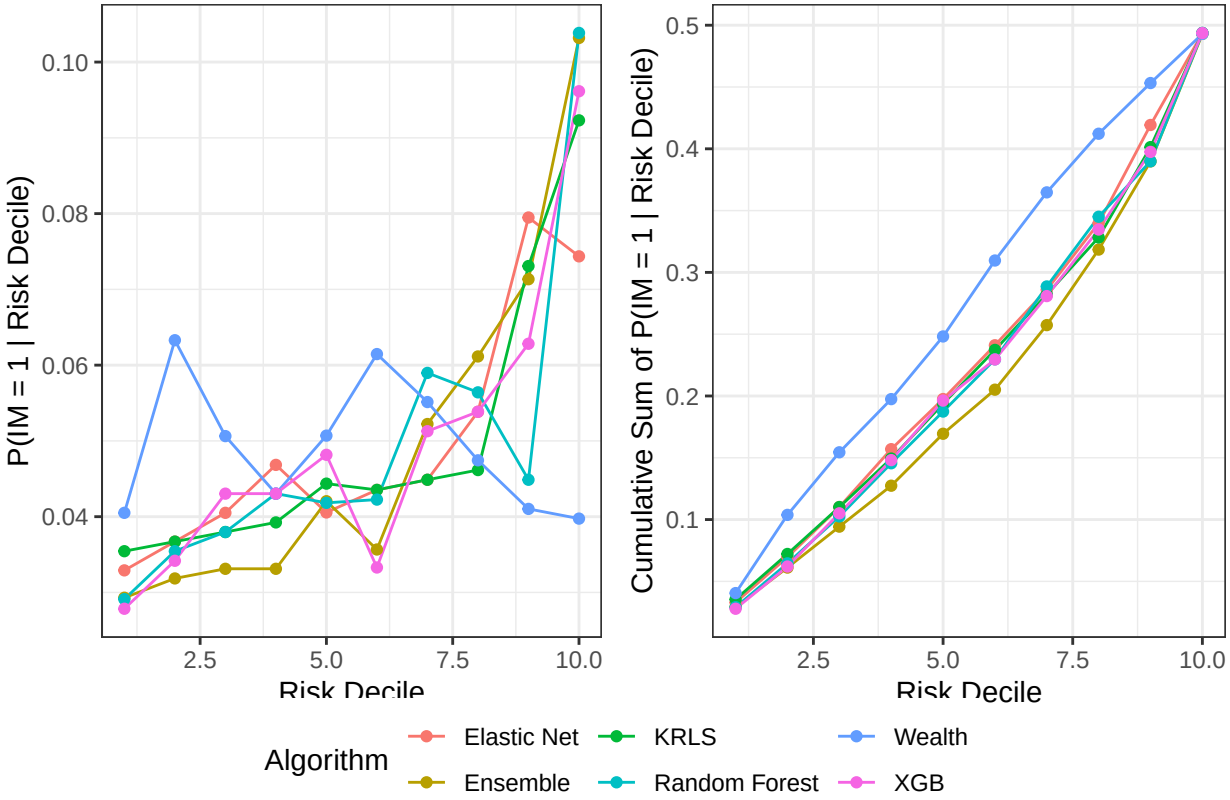

Figure 58: Probability of a mortality given that the observation is in a particular risk decile

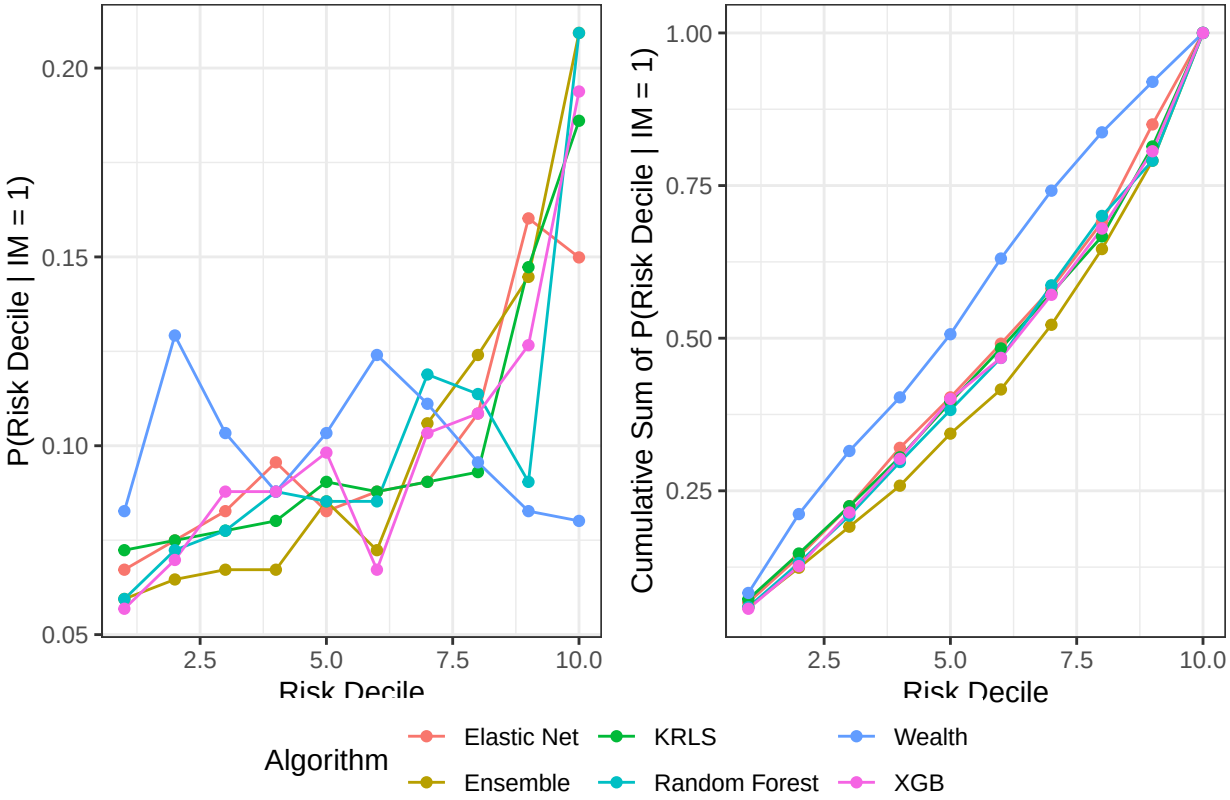

Figure 59: Probability of membership in a particular risk decile given a mortality

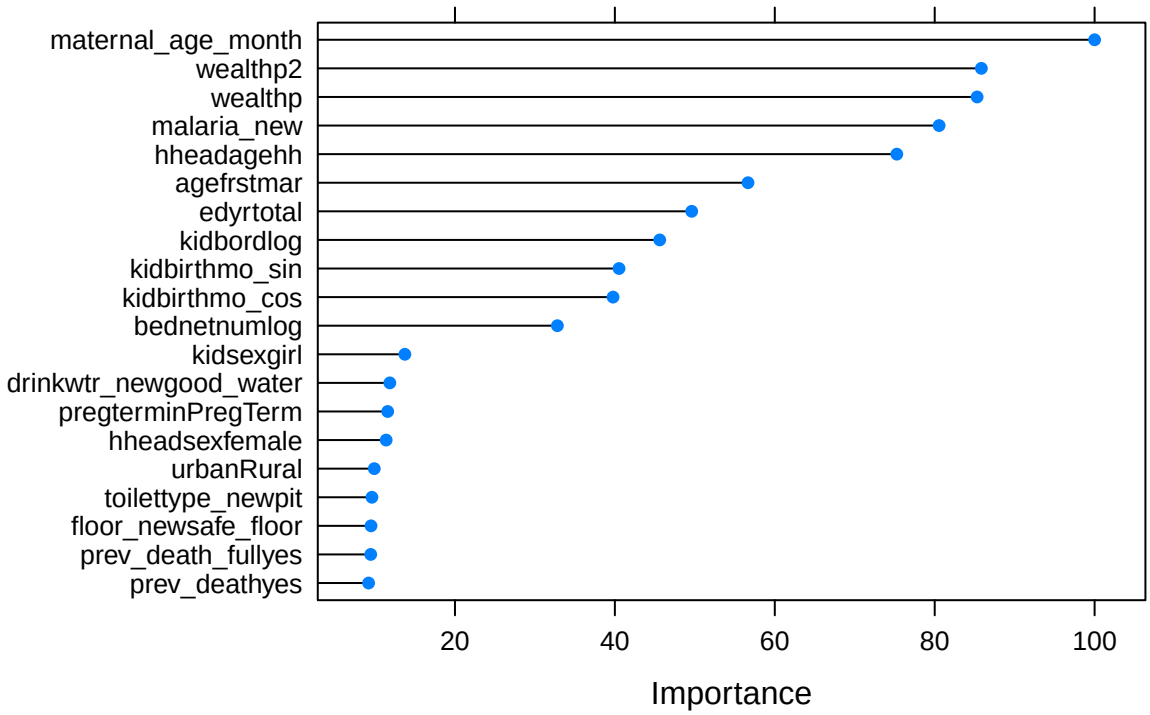

Figure 60: Variable Importance Plot generated by the Random Forest model

B.21.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.21.7 Elastic Net Coefficients

Table 168: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-1.6512	-1.6512
pregterminPregTerm	0.254	0.0817
agefrstmar	-0.0315	-0.1057
hheadagehh	0.0101	0.1065
urbanRural	-0.1266	-0.0607
kidsexgirl	-0.2211	-0.1106
hheadsexfemale	0.1064	0.0401
edyrtotal	-0.0231	-0.0828
drinkwtr_newgood_water	0.0777	0.0385
floor_newsafe_floor	0.0405	0.0188
maternal_age_month	-	-
cookfuel_newclean_fuel	-0.2157	-0.0536
toilettype_newpit	-	-
toilettype_newunimproved	-	-
religion_newChristian	-0.5262	-0.061
religion_newHindu	-	-
religion_newMuslim	0.4863	0.0295
religion_newOther	0.0276	0.0027
bednetnumlog	-0.1409	-0.0733
prev_deathyes	0.2981	0.1149
prev_death_fullyes	0.1422	0.0619
malaria_new	0.055	0.0071
kidbordlog	-0.3687	-0.252
kidbirthmo_sin	-0.1058	-0.0751
kidbirthmo_cos	0.0662	0.0466
wealthp	0.1363	0.0394
wealthp2	-	-

B.22 Zimbabwe

This sample was taken in the year 2015. There are 4630 observations in the dataset, and 3705 observations were used for the training set. Also, 222 observations were removed due to missing values. Finally, there were 236 deaths for the full dataset, and 189 deaths in the training set.

B.22.1 Variables Used

Table 169: Variables included in the model for Zimbabwe

Variable	**Type**
mortality.under12m	categorical
pregtermin	categorical
agefrstmar	numeric
hheadagehh	numeric
urban	categorical
kidsex	categorical
hheadsex	categorical
edyrtotal	numeric
drinkwtr_new	categorical
floor_new	categorical
maternal_age_month	numeric
cookfuel_new	categorical
toilettype_new	categorical
religion_new	categorical
bednetnumlog	numeric
prev_death	categorical
prev_death_full	categorical
kidbordlog	numeric
kidbirthmo_sin	numeric
kidbirthmo_cos	numeric
geo_zw2015	categorical
wealthp	numeric
wealthp2	numeric

B.22.2 Mortality Breakdown by Wealth Quintile

B.22.3 Optimal Parameters

These are the algorithm parameters selected after cross-validation:

Table 170: Optimal Parameters for the Elastic Net algorithm

alpha	**lambda**
1	0.0030947

Table 171: Optimal Parameters for the Random Forest algorithm

mtry
2

Table 172: Optimal Parameters for the XGB algorithm

nrounds	**max_depth**	**eta**	**gamma**	**colsample_bytree**	**min_child_weight**	**subsample**
100	1	0.3	0	0.8	1	1

Table 173: Optimal Parameters for the KRLS algorithm

loss	**epsilon**	**b**	**lambda**
logistic	0.005	66	0.000283

B.22.4 Table of Results

Table 174: Manual Cross-Validation Results

Algorithm	Recall 10	Recall 20	ROC	AUC	MRD	MRR	MSE	F1	Efficiency Gain
Ensemble	0.190	0.333	0.642	0.104	0.010	1.216	0.048	0.929	1.569
KRLS	0.154	0.292	0.600	0.076	0.005	1.108	0.048	0.927	1.265
Elastic Net	0.143	0.308	0.576	0.077	0.005	1.095	0.048	0.926	1.176
Random Forest	0.201	0.338	0.629	0.116	0.019	1.726	0.048	0.930	1.653
Wealth	0.121	0.228	0.570	0.065	2.029	-170.854	89.014	0.925	1.000
XGB	0.175	0.318	0.588	0.081	0.009	1.184	0.048	0.928	1.441

Table 175: Distribution of individuals in the top risk decile of each algorithm among each wealth decile

Wealth Decile	Elastic Net	Ensemble	KRLS	Random Forest	XGB
1	0.274	0.214	0.222	0.201	0.160
2	0.197	0.181	0.197	0.154	0.170
3	0.187	0.168	0.206	0.108	0.173
4	0.149	0.152	0.179	0.100	0.141
5	0.068	0.103	0.092	0.092	0.092
6	0.065	0.092	0.073	0.114	0.106
7	0.022	0.035	0.016	0.073	0.049
8	0.016	0.011	0.008	0.054	0.014
9	0.016	0.038	0.005	0.068	0.076
10	0.005	0.008	0.000	0.035	0.019

B.22.5 Performance Plots

Below are various plots showing the performance for each model considered in Zimbabwe.

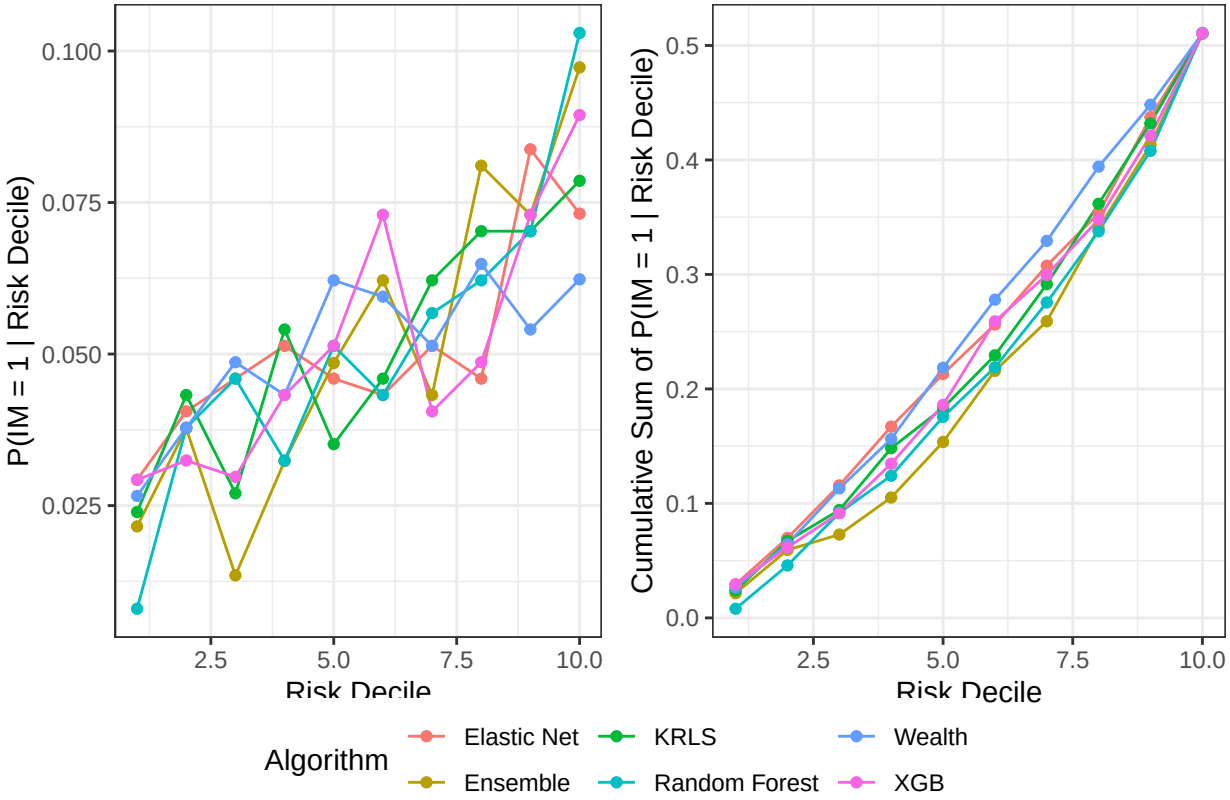

Figure 61: Probability of a mortality given that the observation is in a particular risk decile

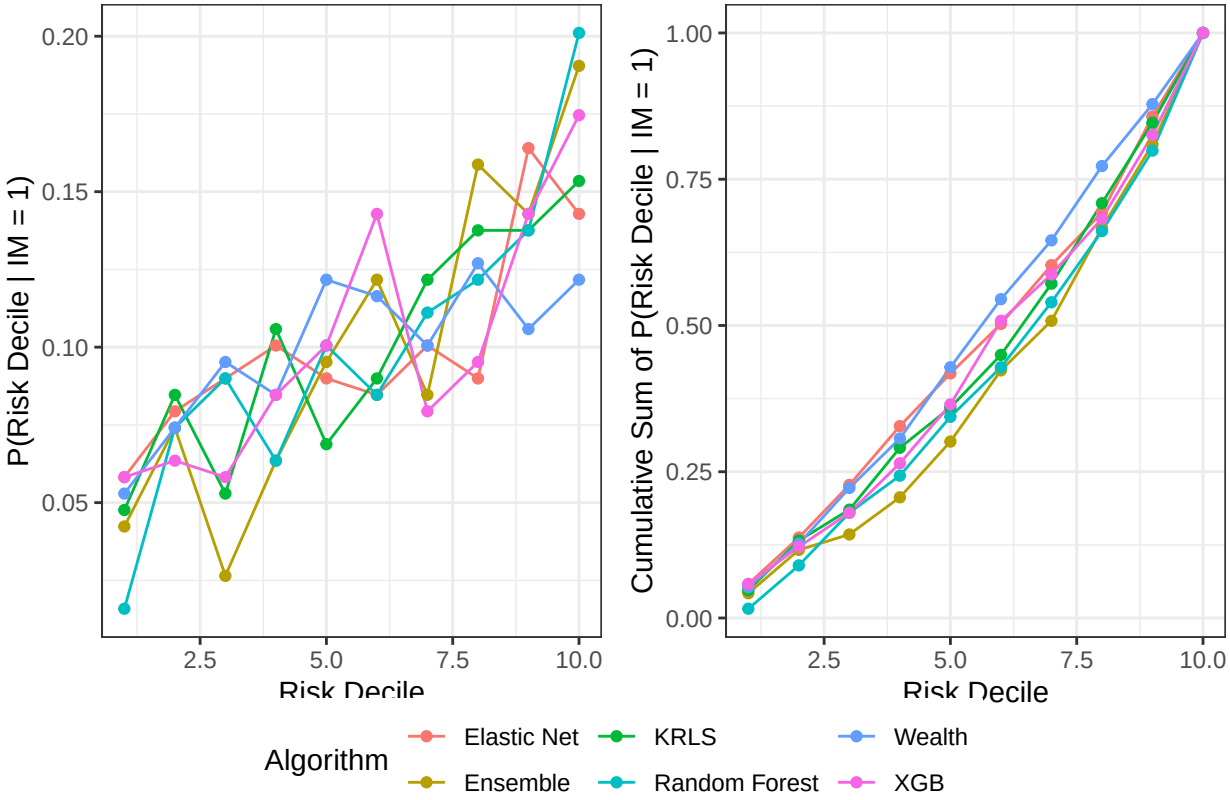

Figure 62: Probability of membership in a particular risk decile given a mortality

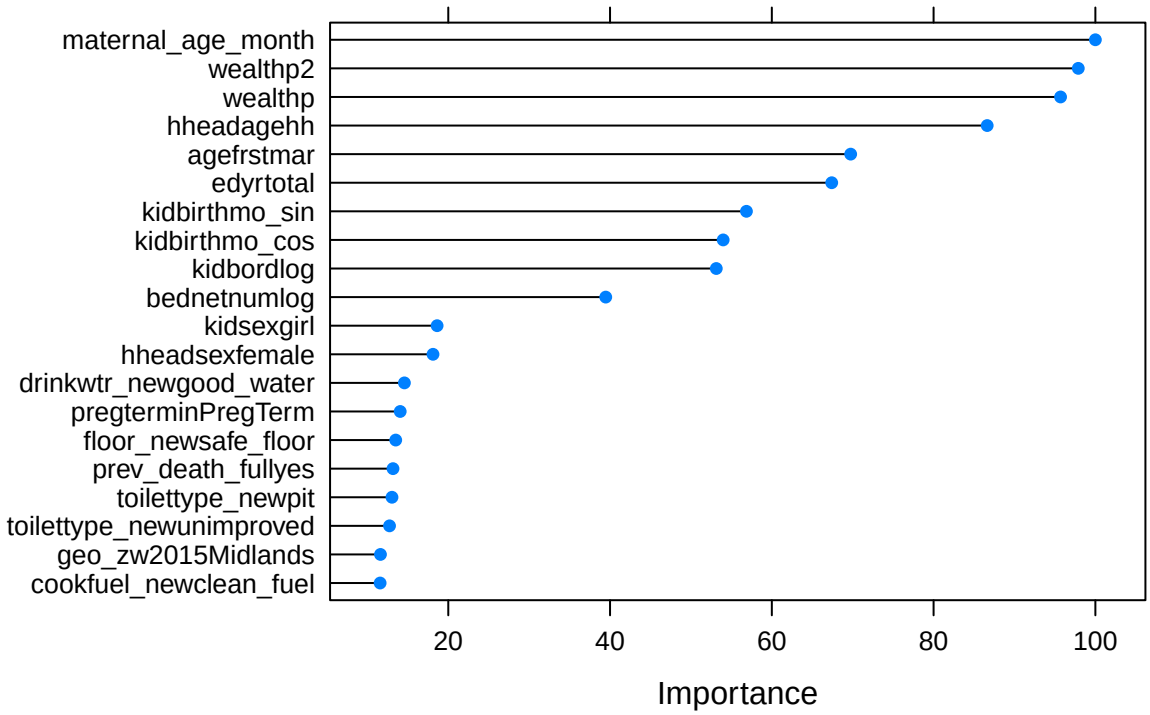

Figure 63: Variable Importance Plot generated by the Random Forest model

B.22.6 Variable Importance Plot

In this section we will show some measures to give some indication of the important variables necessary for predicting infant mortality.

B.22.7 Elastic Net Coefficients

Table 176: Standardized coefficients from the optimal elastic net model after 10-fold cross-validation.

	Coefficients	**Standardized Coefficients**
(Intercept)	-2.1829	-2.1829
pregterminPregTerm	0.0089	0.003
agefrstmar	-0.0021	-0.0076
hheadagehh	-	-
urbanRural	0.0719	0.0349
kidsexgirl	-0.1007	-0.0504
hheadsexfemale	0.1601	0.0768
edyrtotal	-0.0861	-0.2474
drinkwtr_newgood_water	0.0814	0.0349
floor_newsafe_floor	-	-
maternal_age_month	-	-
cookfuel_newclean_fuel	-	-
toilettype_newpit	-	-
toilettype_newunimproved	0.2206	0.097
religion_newChristian	0.151	0.0384
religion_newHindu	-	-
religion_newMuslim	-	-
religion_newOther	-	-
bednetnumlog	-	-
prev_deathyes	0.0259	0.0084
prev_death_fullyes	0.2783	0.0994
kidbordlog	-0.2526	-0.1527
kidbirthmo_sin	0.0561	0.0396
kidbirthmo_cos	-	-
geo_zw2015Mashonaland Central	-0.1399	-0.0452
geo_zw2015Mashonaland East	-	-
geo_zw2015Mashonaland West	-	-
geo_zw2015Matabeleland North	-0.02	-0.0053
geo_zw2015Matabeleland South	-	-
geo_zw2015Midlands	-	-
geo_zw2015Masvingo	-	-
geo_zw2015Harare	-	-
geo_zw2015Bulawayo	-	-
wealthp	-	-
wealthp2	-	-