Grace Saunders, Ravza Aykan, Mariam Seshan
4.30.25

Project Results

Questions

Research Question: This study seeks to explore which measures of economic and societal health are most useful in predicting the receipt of foreign aid. Specifically, we examine the relationship between various economic indicators, such as GDP growth, GNI per capita growth, and poverty headcount ratios, and the amount of Net Official Development Assistance (ODA) received by countries.

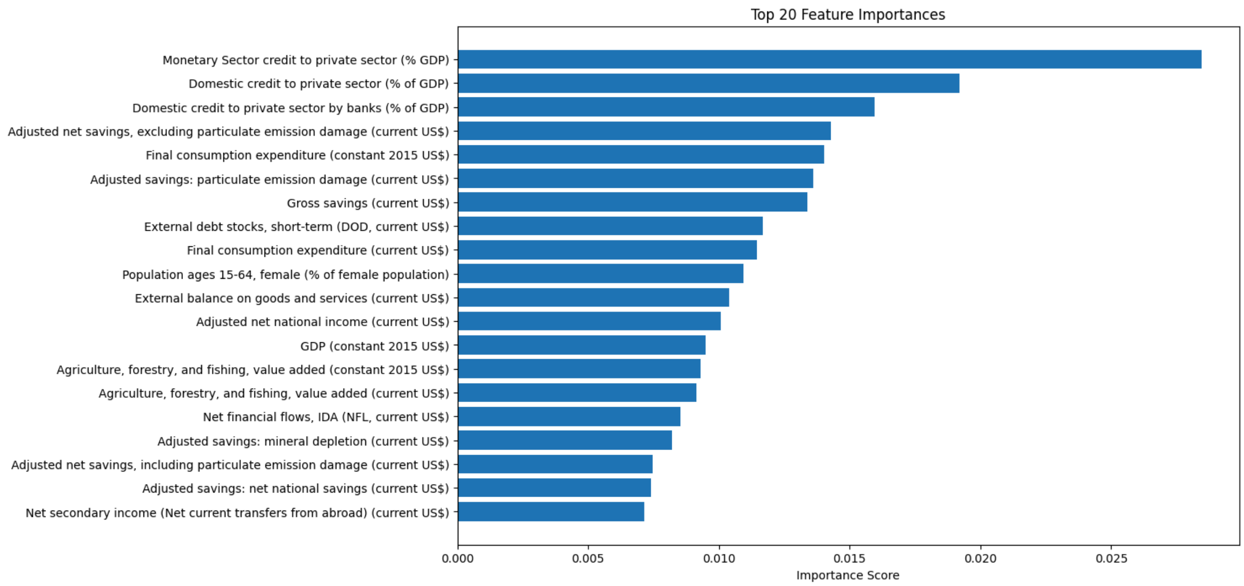
Prediction Question: What drives changes in the amount of official development assistance (ODA) received by a country?

Results

We first constructed a Random Forest model using a broad array of indicators from the World Bank's World Development Indicators dataset, deliberately excluding direct ODA measures to avoid circularity. To create our initial model, we reshaped the data, filtering out indicators which were missing for more than 50% of the country-year observations, and lagged the target variable (Net ODA received in constant 2021 USD) by one year. We then applied an inverse hyperbolic sine transformation to reduce skew. To handle missing values, we used zero imputation and created dummy variables representing whether the values were initially missing. After scaling features with a RobustScaler and splitting the data chronologically (training on years before and including 2017, and testing on years after 2017), we trained a Random Forest Regressor with tuned hyperparameters. The model showed strong in-sample fit, with a training RMSE of 2.59 and an R^2 of 0.76, while the test data yielded an RMSE of 4.04 and an R^2 of 0.53.

We visualized the top 20 most important features to identify which indicators contributed most to the model's predictive power:

Model 1, Top 20 Feature Importances (Github Link)



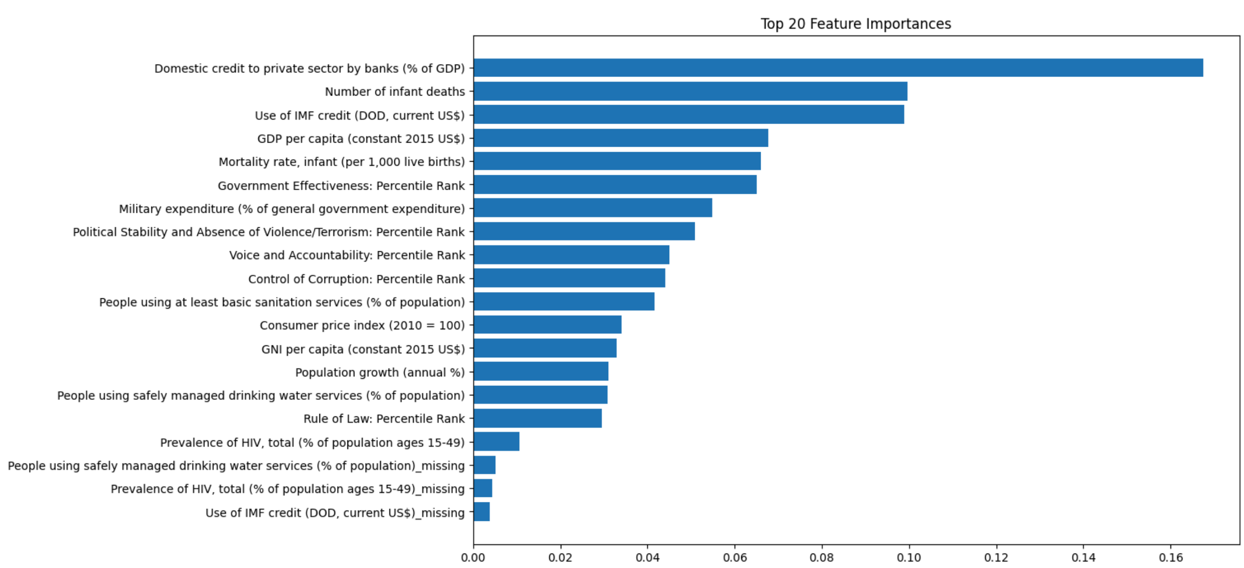
However, we found that this first model included redundant features, such as indicators measured in both constant and current USD or overlapping metrics for similar concepts, which introduced noise and multicollinearity and may have artificially inflated our models performance. Almost all of these indicators were related to, or components of, a country's GDP. To address this, we refined our approach by selecting a targeted subset of 29 distinct and relevant indicators that avoided duplication.

We selected these variables based on relevance to aid allocation, including measures such as GDP per capita, literacy rates, corruption perception, and infant mortality, and made sure that there were no explicitly redundant indicators. We then merged them with the ODA data and filtered the dataset to retain records with adequate coverage in the analysis period. After cleaning, we engineered a one-year lag of the target variable and again transformed it using the inverse hyperbolic sine to reduce skew. We split the data by time, using all observations up to and including 2018 (to account for slightly longer coverage in this refined subset as opposed to using 2017) for training and later years for testing. We standardized the features using a StandardScaler and trained a Random Forest Regressor with the same hyperparameter configuration as in the earlier model.

This second model produced a training RMSE of 2.55 and an R^2 of 0.77, while the test data yielded an RMSE of 4.79 and an R^2 of 0.33. While the model fit the training data well, the drop in performance on the test set suggests lower predictive power of the model. Compared to Model 1, Model 2 performed similarly in training but worse on the unseen test data, suggesting that the broader feature set in Model 1 captured more variation, (though this was likely due to redundancy in features).

We again visualized the top 20 most important features to identify which indicators contributed most to the second model's predictive power.

Model 2, Top 20 Feature Importances (GitHub Link)



We found that, after removing redundant and irrelevant indicators, the top three most important features were domestic credit to the private sector by banks, as a percentage of GDP, the number of infant deaths, and the use of IMF credit (debt outstanding, current USD). The rest of the top 20 features are related to government stability and corruption, and the health of the population.

Summary

Model	Feature Set	Train R^2	Test R^2	Top Feature Domains
1	Broad, partially redundant	0.76	0.53	1. Monetary sector credit to private sector (3 measures) 2. Adjusted net savings (2 measures) 3. Final consumption expenditure
2	Curated, distinct variables	0.77	0.33	1. Domestic credit to private sector by banks 2. Number of infant deaths 3. Use of IMF credit

In both cases, our models suggest that development indicators can help explain ODA patterns, and preprocessing choices, particularly around feature selection, have a meaningful impact on generalization and interpretability.