Executive Summary Qarik Corporate Project Group Bear

Introduction

The World Bank provides loans to developing countries with a view to ending extreme poverty within a generation and boosting shared prosperity. Since the advent of the COVID pandemic, the World Bank has loaned over 157 billion dollars to such countries. This startling statistic shows that in moments of crisis and uncertainty the World Bank is an important source of economic support. On the other hand, we have also found out that about 45% of countries have never gotten a loan from the World Bank when almost 80% of the countries in the world are developing countries. Motivated by the World Bank's desire to make a positive impact in the world, we set ourselves on a mission to understand the underlying factors by which the World Bank distributes their funds by analyzing a corpus of over 3000 unstructured loan documents ranging from 1990 to 2019 provided by Qarik. In this way, we hope to assist developing countries improve their chances of getting their loans approved by the World Bank.

Results

We discovered that the World Bank has a tendency to give larger loans to countries who are going through times of political instability – as opposed to the more risk averse behavior one would expect from traditional Banks. For example, we learned that the World Bank had loaned Iraq and Egypt large amounts of money during (or shortly after) their most trying times – the Iraq War and the Egyptian revolution. Among northern African nations, Egypt received the highest average loan amount 9 times out of the 30 years we examined. Almost all of these years were in the time-span during and after the Egyptian revolution. The average loan amount in the years prior to 2011 was $168.5M, but for 2011 and after, the average loan amount jumped to $546.2M, a massive 224% increase. Which leads us to our first recommendation for countries aspiring to get their loans approved: do not let the internal struggles of your country discourage you from applying for loans, the World Bank may be able to help.

We then worked to quantify how loans have been historically distributed. We were able to successfully find a strong correlation between GDP growth and loan distribution. Countries that have gotten loans have very similar (and relatively low) GDP volatility – a feature we developed ourselves for this project. This finding is compatible with common Bank practices, which uses a country's economics history to determine if they will be able to pay the loan back. This model was evaluated with an ROC AUC score of 0.75 and an accuracy of 70%, indicating there is still room for improvement. We note to our stakeholders that maintaining a consistent GDP growth is correlated with getting a loan, as opposed to having small periods of rapid or artificial GDP growth. This correlation warrants further investigation to determine a causal relationship.

Our most ambitious project was to predict the magnitude of the loan from the features we have extracted for each country. Our best model was a random forest which performed 26% better on RMSE in comparison to our baseline model. We learned that among the features collected, the most impactful features were the sector of the loan, political stability, and GDP per capita – which we quantified using SHAP scores. We note that the classification of loans by sector was done using a Latent Dirichlet Allocation topic analysis model. We recommend to our stakeholders to keep in mind that the sector will have the largest impact in the loan amount – among the different features we considered. In particular, the SHAP analysis indicated that whether a loan focused on the public administration or communication sector was 52% more important than the next most influential variable.

Future Work

Further analysis of new features such as a country's debt sustainability and absorptive capacity are warranted to improve our loan amount predictions – and consequently bring more value to our stakeholders.