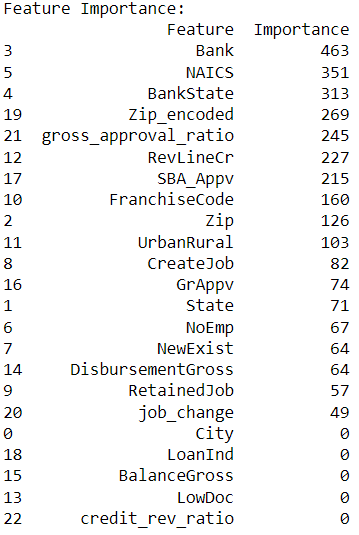
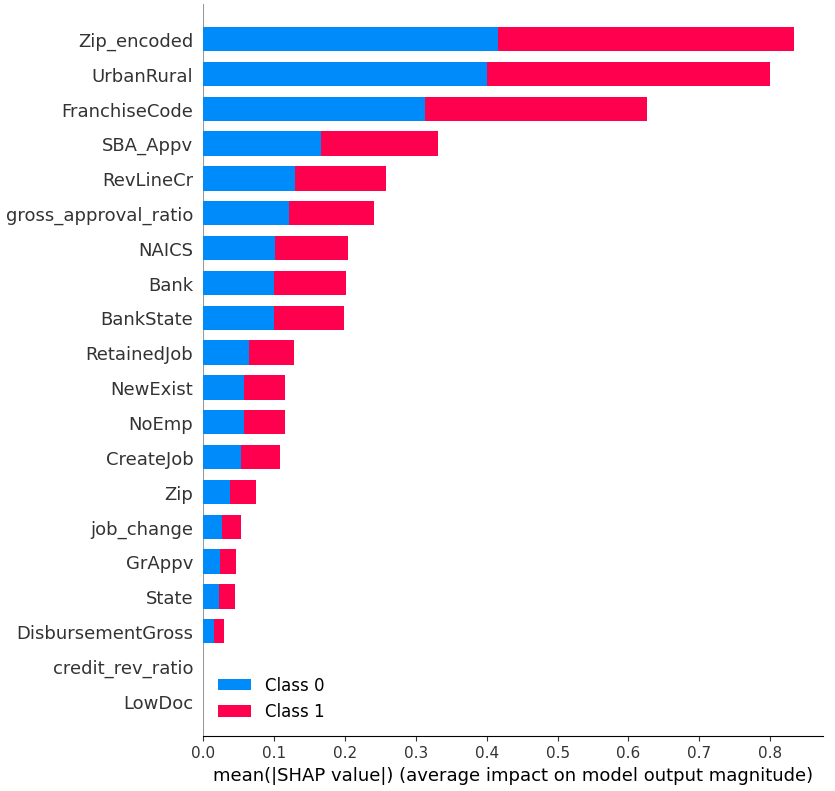
Iori Koh

BUAN 6341

Shap Values/Feature Importance Write Up

These are the Features in ascending order of importance via Shap Values:  


Given this, we can assume that the loan's business and financial characteristics are related to Bank, BankState, and NAICS. These elements include crucial details about the lender, the borrower's location, and the type of business the borrower operates; all of which may have an impact on the borrower's capacity to repay the loan. While NAICS gives information about the borrower's industry and may reflect the borrower's experience and success in that area, Banks and BankState provide information about the lender's location and may indicate the lender's familiarity with the local market. These characteristics may consequently be more indicative of the borrower's capacity to repay the loan and as a result, have higher feature relevance for MIS\_Status.



On the other hand, the SHAP values show that ‘ZIP\_encoded’, ‘UrbanRural’, and ‘FranchiseCode’ are some of the most contributing predictors. Zip code could be a good indicator as area codes and different cities are indicative of wealth. Wealthier cities may impact ‘MIS\_Status’ and the same goes for poor cities; thus the Zip code being a relevant predictor makes sense. In similar logic, UrbanRural would also be a big contributor as rural/urban areas may indicate wealth and credit status. Some features I engineered such as ‘gross\_approval\_ratio’ seem to be a contributing factor also.