Abstract

This is an analysis of South Africa’s (SA) sovereign credit rating (SCR) using Naïve Bayes a Machine learning (ML) technique. Quarterly data from 1999 to 2018 of macroeconomic variables and categorical SCRs were analysed and classified to predict and compare variables used in assigning SCRs. A sovereign credit rating (SCR) is a measurement of a sovereign government’s ability to meet its financial debt obligations. The differences by Credit Rating Agencies (CRA) on rating grades on similar firms and sovereigns have raised questions on which elements truly determine credit ratings. Sovereign ratings were split into two (2) categories that is less stable and more stable. Through data cross-validation for supervised learning, the study compared variables used in assessing sovereign rating by the major rating agencies namely Fitch, Moody’s and Standard and Poor’s. Cross-validation splits the dataset into train set and test set. The research applied cross-validation to reduce the effects of overfitting on the Naïve Bayes Classification model. Naïve Bayes Classification is a Machine learning algorithm, that utilizes the Bayes theorem in classification of objects by following a probabilistic approach. All variables in the data were split in the ratio of 80:20 for the train set and test set respectively. Naïve Bayes managed to classify the given variables using the two SCR categories that is more stable and less stable. Variables classified under more stable indicates that ratings are high or favorable and those for less stable show unfavorable or low ratings. The findings show that CRAs use different macroeconomic variables to assess and assign sovereign ratings. Household debt to disposable income, exchange rates and inflation were the most important variables for estimating and classifying ratings.

Key Words: Sovereign Credit rating, Naïve Bayes, Machine Learning and Macroeconomic variables

JEL Classifications: C51, C52, C53, C58, G17, G24

1. Introduction

Sovereign credit rating (SCR) has become the centre of controversy in the international bond market (Gültekin-Karakaş, Hisarcıklılar and Öztürk, 2011). The move by Credit Rating Agencies (CRA) to give a rating to sovereigns shifted the credit market to another level as investors were put in a dilemma, either to follow published ratings or use their own approaches to rate assets, firms and sovereigns (Osobajo and Akintunde, 2019). Firms, sovereigns and other borrowing institutions were not spared as they tried to find their identity as perceived by markets on creditworthiness. Monitoring the sovereign credit ratings (SCR) give international investors the opportunity to access crucial information on countries that they are planning to invest in (Cantor and Packer, 1994; Ferri et al., 2001; Kabadayi and Çelik, 2015; Mora, 2006). Gogas et al. (2014) claimed that investors, borrowers, issuers and governments use CRAs risk-rating scale obtained on the banks’ ability to meet obligations on debt timely, as provided by CRAs in making investment and financial decisions. Countries with good credit ratings are believed to be financially strong, with stable financial systems and are credit worth (Gültekin-Karakaş et al, 20 11).

A sovereign credit rating (SCR) is a measurement of a sovereign government’s ability to meet its financial debt obligations. After thorough assessments of fundamental information peculiar to a sovereign, CRAs can then identify a rank for that particular sovereign either as capable of redeeming its debts or failing to honour its obligations. Chee et al. (2015) citing De Haan and Amtenbrink (2011) stated that Credit Rating Agencies (CRAs) perform two services namely to provide information and to monitor debtors. Ratings from CRAs faced rejection since the rating industry is oligopolistic. The most prevalent CRAs are Standard & Poor’s (S&P), Moody’s and Fitch Rating Agency, as these three
dominate the market. Bellotti et al. (2011) illustrated that ratings are facing challenges because of opaque methodologies used by rating agencies and failure to predict financial crises like the late 1990 Asian crisis. Bellotti et al. (2011) went on to say that, CRAs do not look ahead on rating assignments but rather look backward. According to Ozturk et al. (2016), the financial crisis of 2008 was a result of high reliance on CRA sovereign ratings. Ratings are genuinely serving a niche financial purpose and generally, local firms are rated lower than the county’s overall sovereign rating (Iyengar, 2010).

Sovereign credit ratings affect interest rates of assets and form other assets’ benchmark indicator for credit risk assessment and so influence the breadth and volume of assets (Ozturk et al., 2016). Ozturk et al. (2016) highlighted that arguments over the accuracy of sovereign ratings augmented the need for more internal credit scoring systems to avoid over dependence on CRAs. Governments are highly interested in credit ratings to improve international capital markets accessibility and reduce borrowing cost. Sovereign ratings are not entirely focused on governments, but for assessing other borrowers of that particular country (Iyengar, 2010). Sovereign rates illustrate the level of default risk associated with a borrowing nation.

Kraussl, (2005) raised that credit rating agencies have substantial influence on the size and volatility of emerging markets lending. Mutize and Nkhalamba (2020) concluded that investors in South Africa’s long-term bonds are more sensitive to negative credit rating events, which are mainly driven by structural problems in the economy. According to Mahomed Karodia and Soni (2014) “the foreign debt rating of South Africa has been over the last several months been continuously downgraded by international rating agencies and, downward growth revisions have become a persistent pattern.” South Africa as an emerging market is highly vulnerable to lending volatility as the country had been facing downgrades since 2009 which affects growth and the financial sector’s stability (Mahomed Karodia and Soni 2014). This study aims to use Naïve Bayes a Machine Learning tool to classify and identify variables used in attaining sovereign credit ratings. Forecasting sovereign ratings assist nations in determining the outlook of their default risk as perceived by those providing funds (Kabadayi and Celik, 2015). Polito and Wickens (2013) and CGFS (2011) highlighted that some scholars accused CRAs of intensifying the debt crisis, increasing costs of borrowing and instigating or exacerbating instability. This problem raises the need for developing and underdeveloped countries to analyse or forecast sovereign ratings. Sovereign ratings failed to clearly identify country specific determinants and impacts of sovereign downgrades. South Africa, as an emerging market with an unpredictable political stability, is vulnerable to negative impacts of sovereign ratings downgrades. Negative expectations or speculating actions towards downgrade can negatively affect the macroeconomic variables leading to financial instability and disturbances in the fiscal or monetary framework.

Previous studies transformed ratings into numerical values losing some information but, in this research, ratings were analysed in their original and categorical format as released by CRAs. The conversion of ratings to numerical values weakened the strength of previous studies data and models in analysing credit ratings. Studies like Pretorius and Botha, (2016), Bhatia, (2002), Afonso, (2003), Kräussl, (2003), Reinhart, (2002), Afonso et al., (2011), Chee et al., (2015), Hill et al., (2010) and Mellios and Paget-Blanc, (2006), Erdem and Varli, (2014), Cantor and Packer, (1994), Ferri et al., (2001), Kabadayi and Çelik, (2015), Mora, (2006) and so on used models that includes panel analysis, probit model, logit, discriminant analysis, multinomial logistic regression and other statistical models which assume linear relationships and normal distribution lacking on the ability to perfectly model sovereign ratings. Sovereign ratings are determined by various factors that may have either a linear, non-linear relationship, normally distributed and non-normal. Therefore, the study to applied Naïve Bayes which classifies variables and is not affected by the state of explanatory variables being normally distributed or having a linear relationship.

Studies by Cantor and Packer, (1994), Ferri et al., (2001), Kabadayi and Çelik, (2015), Mora, (2006) Pretorius and Botha, (2016), Bhatia, (2002), Afonso, (2003), Kräussl, (2003), Reinhart, (2002), Afonso et al., (2011), Chee et al., (2015), Hill et al., (2010) and Mellios and Paget-Blanc, (2006), Erdem and Varli, (2014) and so on which conducted an approach of grouping or analysing cross country sovereign ratings failed to clearly identify country specific determinants and impacts of sovereign ratings. Pulling countries together fails to consider the differences in development between the countries. The study aimed to apply a new model that classifies variables used to predict SCRs, to assist governments prevent downgrades and promote upgrades restoring financial stability. The research further examined sovereign rating links with macroeconomic indicators to exhume the determinants of the sovereign credit ratings from Fitch, Moody’s and Standard & Poor’s (S&P). Naïve Bayes was chosen as it could improve the analysis of sovereign ratings far better than statistical models as it was applied in comparing variables applied by Fitch, Moody’s and Standard & Poor’s (S&P) in ratings sovereigns through classification.
This paper is structured as follows: Section 2 provides a discussion of empirical and theoretical literature. Section 3 illustrates the machine learning tool which is Naïve Bayes classification methodology applied to classify sovereign ratings. Section 4 presents empirical results and section 5 discusses the conclusion.

2. Literature Review

The literature review has three parts which includes theoretical literature, empirical literature and analysis of literature.

2.1. Theoretical literature on determinants of SCR

Chee et al (2015) highlighted that the political and inflation theory asserts that inflation is a measurement of incumbent government’s self-restrain in not being opportunist for election benefits as unhappy may increase the possibilities of political instability. CRAs use rating forms to assess sovereign debt creditworthiness or international borrowing by countries. CRAs use various factors like social, economic and political aspects. Inflation and business cycles are linked to the political aspect as this theory notes that people use voting to illustrate their dissatisfaction with high inflation and negative business. Political instability or dissatisfaction by citizens affects governance and reduces the ability of a country to raise income to honor debt obligations. Once political instability and dissatisfaction from citizens kicks in, lenders are skeptical on providing funds to the country as it is deemed not creditworthy. The economic aspect is indicated by quantifiable macroeconomic factors that outline the capacity of a country to service its external debt. The ability of a government to plan and manage debt efficiently is indeed a big factor as failure will lead to high risk of default. Chee et al (2015:43) argued that “though a country’s capacity in repaying external debts is unable to be quantified; however, the sovereign creditworthiness can be measured by observing government’s sovereign debt management behavior, which could be traced by the monetary policy and fiscal policy.”

2.2. Empirical Literature Review

The differences by CRA on rating grades on similar firms and sovereigns have raised questions on which elements truly determine credit ratings. Ozturk et al. (2016) citing Cantor and Packer (1996), and Ferri et al. (1999) highlighted that early literature examined to what extent the variation in sovereign ratings could be explained by country specific variables and identified certain macroeconomic, financial and political variables that bode well in explaining the variations in sovereign credit ratings. Kabadayi and Çelik (2015) in reviewing literature highlighted that Archer et al. (2007), Butler and Fauver (2006), Cantor and Packer (1996), Ferri et al. (1999), Mora (2006) and Ratha et al. (2011) applied sovereign ratings as quantitative dependent variables. Ferri et al. (2001), Gaillard (2006), Hill et al. (2010), Hu et al. (2002) and Mora (2006) used sovereign ratings as qualitative dependent variables.

Afonso et al. (2011) conducted a study from 1995 to 2005 on the determinants of sovereign debt ratings using rating notations from the three-leading international CRAs. The study made use of a linear regression method and a random effect ordered probit response model. They applied a unique process that distinguished short term and long-term effects from several macroeconomic and fiscal explanatory variables on a country’s sovereign debt rating. According to Afonso et al. (2011) changes in GDP per capita, GDP growth, government debt, and government balance had a short-run impact on a country’s credit rating, whereas government effectiveness, external debt, foreign reserves and default history were important long-run determinants. Novotná (2012) applied and compared three selected approaches for credit rating prediction, linear discriminant analysis, and logistic regression as well as Naive Bayes based on data of European companies’ financial indicators. Novotná (2012) concluded that discriminant analysis and logistic regression models are relatively simple to use and achieve high classification ability. Various contradicting models were applied in assessing sovereign ratings determinants producing conflicting results.

2.3. Summary of Literature reviewed

The literature review illustrates that most of the studies conducted in analysing the determinants and impacts of sovereign credit ratings applied cross-sectional or cross-country approach. The weakness of this approach is that grouping countries fails to capture country specific conditions as countries are not similar or some are well developed than others like in a case of South Africa and China in the BRICS countries making findings unreliable. Unique models applied by different researchers produced
conflicting results. This study applies a country specific approach looking at the determinants of sovereign ratings for South Africa using the country's specific economic indicators through a Naïve Bayes Classification model.

3. Research Methodology

This study used a Machine Learning (ML) algorithm that is Naïve Bayes Classifier to classify and analyse SCR from each one of the major CRAs namely Fitch, Moodys and S&P using micro and macroeconomic indicators. Machine learning tools are representing an ever-increasing research area (Azeem, Palomba, Shi & Wang, 2019). Machine learning is a branch of artificial intelligence that aims at enabling machines to perform their jobs skilfully by using intelligent software and the statistical learning methods constitute the backbone of intelligent software that is used to develop machine intelligence (Mohammed, Khan & Bashier, 2016:4). Thus, they are techniques that utilise statistical methods in determining patterns in massive data. The Naïve Bayes Classification ML model captured the relationship between SCR and economic indicators by clearly identifying crucial determinants of SCR. This study focused on studying if sovereign credit rating is based on macroeconomic variables and if CRAs use the same variables to model sovereign credit ratings. SCR were split into two (2) categories that is less stable and more stable. Less stable indicating low sovereign credit ratings or junk status borrowers and more stable representing higher sovereign ratings with high creditworthiness.

According to Chee et al, (2015:44) citing Feder and Uy (1985) and Cantor and Packer (1996) modelled sovereign credit rating using variables that impact a country's international creditworthiness and the reduced formulae is as follows:

\[ CCCW = f( \text{economic indicators associated to debt repaying capacity of debtor countries} ) \]  

The following function illustrates the relationship between sovereign ratings and economic variables:

\[ SR_{jit} = f(Macro \text{ and } Micro \text{ Economic Variables}) \]  

Where ‘j’ represents the specific CRA between Fitch, Moody and S&P, ‘t’ represent time in quarterly from 1999 to 2018 and ‘i’ represent a dummy for ‘more stable rating’ or ‘less stable rating’. On this research the focus was only on South Africa. The model specification is as follows:

\[ SCR_{jit} = \beta_0 + \text{REER} \beta_1 + \text{PIR} \beta_2 + \text{HDDI} \beta_3 + \text{UR} \beta_4 + \text{GDPpc} \beta_5 + \text{CPIH} \beta_6 + \text{FDGDP} \beta_7 + \text{BOP} \beta_8 + \text{CAB} \beta_9 + \epsilon \]  

3.1. Cross-Validation

Cross-validation is applied in Machine learning tools to solve the problem of overfitting. Cross- validation splits the dataset into train set and test set. The train set is used to train the model to capture trends or patterns in the dataset and then independent variables in the test set are used to predict values of the dependent values in the test set. The research applied cross-validation to reduce the effects of overfitting on the Naïve Bayes Classification model. All variables in the data were split in the ratio of 80:20 for the train set and test set respectively. The ML models were trained using 80% of the data and then used 20% of the explanatory variables in the test set to estimate 20% of the response variable in the test set.

The financial and economic indicators from South Africa (SA) were compiled into quarterly time series data from 1999 to 2018 giving a total of 80 observations. The data was taken from five years after the time South Africa gained independence up to 2018 since 2019 values were not readily available. The data were collected from various sources like Quantec Easy data, Statistics South Africa (Stats SA), Trading Economics, Thomson Reuters and the South African Reserve Bank (SARB). The data sources have been known to have reliable and valid data of these variables. The figure below shows definition of variables, data sources and priori expectations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition and proxies</th>
<th>Data source, format and transformation</th>
<th>A priori expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sovereign ratings</td>
<td>Sovereign rating is an independent assessment of the creditworthiness of a country or sovereign entity. Sovereign credit ratings can give investors insights into the level of</td>
<td>Ratings from the three major CRAs. Collected in ordinal or categorical format and were transformed to numerical values using a scale of 1 to 21. The 21 denotes the highest rating and a drop is a downgrade whilst a 1 the lowest rating denotes default. A rise from 1 is an upgrade. Downgrades are</td>
<td>21 denotes the highest rating and a drop is a downgrade whilst a 1 the lowest rating denotes default. A rise from 1 is an upgrade. Downgrades are</td>
</tr>
</tbody>
</table>
risk associated with investing in the
debt of a particular country,
including any political risk.

highest rating is denoted by 21
whilst the lowest rating is
represented by 1. Kabadayi and
Çelik (2015)

Trading Economics Website
Principal Component Analysis
(PCA)

unfavourable as they are
associated with high borrowing
costs and less access to funds
and vice versa.

### Table 1: Summary of Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Sources</th>
<th>Favourability</th>
</tr>
</thead>
<tbody>
<tr>
<td>REER,</td>
<td>Real Effective Exchange rates</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when increase – REER,</td>
</tr>
<tr>
<td>PIR,</td>
<td>Prime Interest Rates</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – PIR,</td>
</tr>
<tr>
<td>HDDI,</td>
<td>Housed Hold Debt to Disposable Income</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – HDDI</td>
</tr>
<tr>
<td>UR,</td>
<td>Unemployment rate</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – UR,</td>
</tr>
<tr>
<td>GDPpc,</td>
<td>Gross Domestic Product (GDP) percentage change</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when increase – positive GDPpc,</td>
</tr>
<tr>
<td>CPIH</td>
<td>Consumer price index Headline</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – CPIH</td>
</tr>
<tr>
<td>FDGDP,</td>
<td>Foreign Debt to GDP,</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – FDGDP</td>
</tr>
<tr>
<td>BOP,</td>
<td>Balance of Payments</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – deficit BOP</td>
</tr>
<tr>
<td>CAB</td>
<td>Current Account Balance</td>
<td>SARB and Quantec Easy Data databases.</td>
<td>Favourable when decreasing – deficit CAB</td>
</tr>
</tbody>
</table>

Source: By Author

### Figure 1. Summary of variables

### 3.2 Naïve Bayes Classification

Naïve Bayes Classification is a Machine learning algorithm, that utilizes the Bayes theorem in classification of objects by following a probabilistic approach. Naïve Bayes is based on the Bayes theorem also known as the Bayes Rule. The Bayes Rule used in Naïve Bayes is derived from two notations which are

\[
P(A/B) = \frac{P(A \text{ and } B)}{P(B)} \quad [4]
\]

and

\[
P(B/A) = \frac{P(A \text{ and } B)}{P(A)} \quad [5]
\]

where

- \( P(A) \) is the probability of \( A \) occurring, \( P(B) \) is the probability of \( B \) occurring, \( P(A \text{ and } B) \) is the joint probability of \( A \) and \( B \), that is, the event that both \( A \) and \( B \) occurs,

- \( P(A/B) \) is the conditional probability of event \( A \) occurring given that \( B \) has occurred and similiar \( P(B/A) \) is the conditional probability of event \( B \) occurring given that \( A \) has occurred (Keller, 2018; Berrar, 2018).

Looking at equations 1 and 2, it can be noted that

\[
P(A \text{ and } B) = P(A/B)P(B) = P(B/A)P(A) \quad [6]
\]

The terminologies of the Bayesian theorem are

- \( A \) is known as the proposition and \( B \) is the evidence
- \( P(A) \) represents the prior probability of the proposition
• $P(B)$ represents the prior probability of evidence

• $P(A|B)$ is called the posterior

• $P(B|A)$ is the likelihood

Thus, the Bayes theorem can be summed up as:

$$Posterior = (Likelihood) \cdot (Proposition\ prior\ probability) / Evidence\ prior\ probability$$  \hspace{1cm} [7]

The Bayes theorem is used to calculate the conditional probability, which is the probability of an event occurring based on information about the events that have occurred in the past (He, Zhang & Shi, 2012; Keller, 2018; Berrar, 2018). The above equations are of single predictors but in the real world, there are more than one predictor variables and for a classification problem, there is more than one output class.

### 3.3. Naïve Bayes Algorithm

The Naïve Bayes algorithm is a classification algorithm based on the Bayes rule which assumes that the attributes $X_1, X_2, \ldots, X_n$ are conditional independent of one another, given $Y$ (Krichene, 2017). According to Mitchell (2010) the value of this assumption simplifies the representation of $P(X/Y)$ and the problem of estimating it from the training data. The objective of a Naïve Bayes algorithm is to measure the conditional probability of an event with a feature vector $X_1, X_2, \ldots, X_n$ belonging to a particular class $Y_i$, where

$$P(Y_i|x_1, x_2, \ldots, x_n) = \frac{P(x_1, x_2, \ldots, x_n|Y_i) \cdot P(Y_i)}{P(x_1, x_2, \ldots, x_n)}$$  \hspace{1cm} [8]

Computing the above equation, one obtain:

$$P(x_1, x_2, \ldots, x_n|Y_i) \cdot P(Y_i) = P(x_1, x_2, \ldots, x_n|Y_i) \cdot P(Y_1) \cdot P(Y_2) \cdot \ldots \cdot P(Y_k)$$

$$= P(x_1|X_2, \ldots, X_n, Y_i) \cdot P(x_2|X_3, \ldots, X_n, Y_i) \cdot \ldots \cdot P(x_n|Y_i) \cdot P(Y_i)$$  \hspace{1cm} [9]

The conditional probability, i.e., $P(x_i|x_{i+1}, \ldots, x_n, Y_i)$ adds up to $P(x_i|Y_i)$ since each predictor variable is independent in Naïve Bayes. The final equation becomes

The final equation comes down to:

$$P(x_1, x_2, \ldots, x_n|Y_i) = \prod_{j=1}^{n} P(x_j|Y_i) \cdot \frac{P(Y_i)}{P(x_1, x_2, \ldots, x_n)}$$ \hspace{1cm} [10]

Here $P(x_1, x_2, \ldots, x_n)$ is constant for all the classes, therefore one gets:

$$P(Y_i|x_2, x_3, \ldots, x_n) \propto \prod_{j=1}^{n} P(x_j|Y_i) \cdot P(Y_i)$$ \hspace{1cm} [11]

### 3.4. Steps used in Naïve Bayes Model

The following steps were used in the development of the Naïve Bayes Model:

**Step 1: Obtaining the data and formatting**

At this stage data was imported into R and the variables were formatted in such a way that the continuous variable in the data frame were interval (FDGDP, REER, PIR, BOB, CAB, UR, GDPpc, HDDI, and CPIH) and the categorical variable or response variable, Outcome was indicated as a factor variable.
Step 2: Visualization of the data
For developing a Naïve Bayes classification model, one need to make sure that the independent variables are not highly correlated. The command `Pairs, panel` command was used to determine how the variables were correlated and in this case correlations and histograms were produced. Box plots of the data were also constructed using the command `ggplot`. The boxplots were able to indicate whether the category “less stable” had values higher or less than the category “more stable”. However, if one of the averages was higher than the other then there was some potential of developing a classification model which was the case in all variables. A significant amount of overlap between some of the boxplots showed that the model was unlikely to be 100% accurate due to the overlap.

Step 3: Data partitioning
The random seed `set.seed` of 1234 command was used to make the analysis repeatable and the data was split into 80% training data and 20% test data.

Step 4: The Naïve Bayes model
The Naïve Bayes model was fitted to the data by using the function `naïve_bayes` and the model was plotted using the function `plot(model)`. The plotted model showed the density plot for each continuous variable showing the density of those observations which were less stable and those which were more stable and where they overlap.

Step 5: Prediction
The predictions were stored in a matrix called `p` and the type of prediction was probability. Predictions were done for the train data and the `head(cbind(p, train))` command was used to show the probabilities for each observations. The result showed each of the 80 observations the chances of being less stable or more stable such that if probability was higher than 0.5 for that category it meant that the observation was likely to fall into that category. For example, an 0.64 and 0.36 meant that the observation was likely to fall into less stable, that is a 64% chance which was category 0.

Step 6: Construction of confusion matrix
Predictions for the train data were stored in `p1` and for the test data and in `p2`. Confusion matrixes were developed for the train data and the test data by using the command `table`. Misclassification were also obtained for both data sets.

Step 7: Naive Bayes model with Kernel
To improve the misclassification rates, the Naïve Bayes model was developed by using Kernel. Kernel based densities tend to perform better when numerical variables are not normally distributed. The model was also plotted, and confusion matrixes developed for both train and test data. This showed an improvement on the misclassification as they become less than the first classifications without Kernel density.

4. Presentation of Empirical Results

The empirical results are illustrated in five parts.

4.1. Naïve Bayes Model

Naïve Bayes Classifier is a Machine learning tool used to make predictions. The purpose was to use a classification algorithm which assists in building a fast Machine learning model that can make quick prediction or identification of variables utilized by CRAs. Naïve Bayes was used as part of the Machine learning technique to determine how the micro and macroeconomic indicators FDGDP, REER, PIR, BOP, CAB, UR, GDPpc, HDDI and CPIH could assist in classing the sovereign ratings into less stable and more stable. The data was divided into 80% training and 20% test data and the coding was “1” for more stable and “0” for less stable. The Naïve Bayes was performed using the R package. A Naïve Bayes Classification model requires that the independent variables are not highly correlated. The correlation matrix is shown in Figure 2.
The results show that no independent variable have high correlations which are more than .7 indicating that there is no multicollinearity. The results are presented in the next subsections for Fitch, Moody’s and Standard & Poor’s (SNPoor).

4.2. Naïve Bayes Model for Fitch

Before the Naïve Bayes model could be fitted visualization of data was done using boxplots. The boxplots are shown in Figure 3.
The left side boxplot is the one for less stable whereas the right side is the one for more stable. A significance amount of overlap implies that the model is not likely to be 100% accurate. There is some potential of developing a classification model since for all the variables the values show different medians for less stable and for more stable. The variables REER, GDPpc, HDDI and CPIH showed that high values favored the rating being more stable whereas for the variables PIR, BOP, CAB, UR and FDGDP high values were associated with those less stable. The variables BOP and HDDI showed not much overlap between the less stable and the more stable suggesting that there will not be a lot of misclassification on these variables. CPIH had a lot of overlap between less stable and more stable such that there is potential of classification for the variable.

A Naïve Bayes model was then fitted and the prior probabilities were obtained in Figure 4.

<table>
<thead>
<tr>
<th>Probabilities</th>
<th>Less stable</th>
<th>More stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priori probabilities</td>
<td>0.3768</td>
<td>0.6232</td>
</tr>
</tbody>
</table>

A Naïve Bayes model was then fitted and the prior probabilities were obtained in Figure 4.

The results show that for the Fitch data the probability that a quarterly index is less stable was .38 and .62 for more stable. For the Fitch data most of the indices fall under more stable as evidenced by a prior probability of more than .5. The means and standard deviations for each independent variable are shown in Figure 5.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Less stable</th>
<th>More stable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>REER</td>
<td>82.7255</td>
<td>6.7467</td>
</tr>
<tr>
<td>PIR</td>
<td>13.1186</td>
<td>3.1538</td>
</tr>
<tr>
<td>BOP</td>
<td>33442.88</td>
<td>33921.41</td>
</tr>
</tbody>
</table>
Figure 5. Summary statistics by outcome for Fitch Naïve Bayes Model

Looking at Figure 5, it can be noted that the variables REER, GDPpc and HDDI have high means when the index is more stable where as PIR, BOP, CAB, UR, CPIH and FDGDP have lower mean values when the index is more stable. The findings agree with the analysis done by Mellios and Paget-Blanc (2006 and Chee et al 2015) and concluded that exchange rates are one of the variables used by rating agencies to measure a country’s creditworthiness. The results are consistent with findings by Bissoondoyal - Bheenick et al (2005), Mellios and Paget-Blanc (2006), Iyengar (2010), Afonso, Gomes and Rother (2011), Arefevs, Giits Braslinš (2013), Sánchez-Monedero et al (2014), Kabadayi and Çelik (2015), Ivanovic et al (2015), Chee et al (2015), De Moor et al (2018) and Cantor and Packer (1996) that GDP growth is a crucial variable in determining sovereign credit ratings.

The confusion matrix for the train and test data are shown in Figure 6.

For the train data the true positive for class 1 = “More stable” was 39 whereas false positive was 1, for class 0 = “Less stable”, true positives was 25 and false positive was 4 for the train data. The correctly classified observations for the train data was 92.75% whereas the percentage misclassified was 7.25%. For the test data, all the observations were properly classified. This was a good fit since percentage of misclassification is low.

When the model with Kernel smoothing was done the confusion matrix is shown in Figure 7.

For the train data the true positive for class 1 = “More stable” was 40 whereas false positive was 1, for class 0 = “Less stable”, true positives was 25 and false positive was 3 for the train data. The correctly classified observations for the train data was 94.2% whereas the percentage misclassified was 5.8%. For the test data, all the observations were properly classified. Use of the Kernel smoothing increased the classification rate by almost 2%. The error rate using Fitch data was less than 6%.
4.3. Naïve Bayes Model for Moodys

The boxplots obtained of the distribution of variables by outcome are shown in Figure 8.
A classification model can be developed since for all the variables the values show different medians for less stable and for more stable. The variables REER, GDPpc and HDDI showed that high values favored the rating being more stable whereas for the variables PIR, BOP, CAB UR, CPIH and FDGDP high values were associated with those less stable. The variables BOP, UR and HDDI showed not much overlap between the less stable and the more stable suggesting that there will not be a lot of misclassification on these variables. However, the variable CPIH showed a lot of overlap indicating that there is potential of a lot of observations being misclassified.

A Naïve Bayes model was then fitted, and the prior probabilities were obtained in Figure 9.

<table>
<thead>
<tr>
<th>Probabilities</th>
<th>Less stable</th>
<th>More stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priori probabilities</td>
<td>0.5362</td>
<td>0.4638</td>
</tr>
</tbody>
</table>

The probability that a quarterly index is less stable was .54 and .46 for more stable for the Moody's data. It can be concluded that for the Moody's data most of the quarterly indices fall under less stable. The means and standard deviations for each independent variable are shown in Figure 10.
The independent variables with high means when the index is more stable are REER, GDPpc and HDDI. Variables such as PIR, BOP, CAB, UR, CPIH and FDGDP have lower mean values when the index is more stable. These findings agree with previous analyses by Mellios and Paget-Blanc (2006) and Chee et al. (2015), indicating that exchange rates are a key factor in measuring a country’s creditworthiness. The results are consistent with earlier studies by Bissoondoyal-Bheenick et al. (2005), Mellios and Paget-Blanc (2006), Iyengar (2010), Afonso, Gomes and Rother (2011), Arefjevs, Ģirts Brasliņš (2013), Sánchez-Monedero et al. (2014), Kabadayi and Çelik (2015), Ivanovic et al. (2015), Chee et al. (2015), De Moor et al. (2018) and Cantor and Packer (1996), which emphasized the importance of GDP growth in determining sovereign credit ratings.

CPIH shows less difference between means for less stable and more stable levels.

The confusion matrix for the train and test data is shown in Figure 11.

For the train data, the true positive for class 1 = “More stable” was 31 whereas false positive was 3, for class 0 = “Less stable”, true positives was 34 and false positive was 1. The correctly classified observations for the train data was 94.2% whereas the percentage misclassified was 5.8%. For the test data, the percentage correctly classified was 90.91% whereas those misclassified was 9.09%. The fit was better for the train data than for the test data.

When the model with Kernel smoothing was fitted to the Moodys data, the confusion matrix is shown in Figure 12.
For the train data the true positive for class 1 = “More stable” was 32 whereas false positive was 3, for class 0 = “Less stable”, true positives was 33 and false positive was 0 for the train data. The correctly classified observations for the train data were 94.2% whereas the percentage misclassified was 5.8%. For the test data, the percentage correctly classified was 90.91% whereas those misclassified was 9.09%. The fit was better for the train data than for the test data. The error rate using Moodys data was less than 6% which is good.

4.4. Naive Bayes Model for SNPoor

The boxplots obtained for the distribution of variables by outcome are shown in Figure 13.
A classification model can be developed since for all the variables the values show different medians for less stable and for more stable. The variables REER, GDPpc and HDDI showed that high values favored the rating being more stable whereas for the variables PIR, BOP, CAB UR, CPIH and FDGDP high values were associated with those less stable. The variables REER and BOP showed not much overlap between the less stable and the more stable suggesting that there will not be a lot of
misclassification on these variables. However, the variable CAB and CPIH showed a lot of overlap indicating that there is potential of a lot of observations being misclassified.

A Naïve Bayes model was then fitted, and the prior probabilities were obtained in Figure 14.

<table>
<thead>
<tr>
<th>Probabilities</th>
<th>Less stable</th>
<th>More stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priori probabilities</td>
<td>0.4638</td>
<td>0.5362</td>
</tr>
</tbody>
</table>

Source: By Author

**Figure 14. Prior probabilities for the SNPoors data**

The probability that a quarterly index is less stable was .46 and .54 for more stable for the SNPoors data. It can be concluded that for the SNPoors data most of the quarterly indices fall under more stable. The means and standard deviations for each independent variable are shown in Figure 15.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Less stable</th>
<th>More stable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>REER</td>
<td>82.3832</td>
<td>6.1223</td>
</tr>
<tr>
<td>PIR</td>
<td>12.3932</td>
<td>3.2216</td>
</tr>
<tr>
<td>BOP</td>
<td>18533.7812</td>
<td>45731.1961</td>
</tr>
<tr>
<td>CAB</td>
<td>-22943.47</td>
<td>25261.66</td>
</tr>
<tr>
<td>UR</td>
<td>25.8438</td>
<td>1.6594</td>
</tr>
<tr>
<td>GDPpc</td>
<td>1.9906</td>
<td>1.3684</td>
</tr>
<tr>
<td>HDDI</td>
<td>65.6906</td>
<td>10.3236</td>
</tr>
<tr>
<td>CPIH</td>
<td>74.6585</td>
<td>27.8133</td>
</tr>
<tr>
<td>FDGDP</td>
<td>13.1156</td>
<td>3.8203</td>
</tr>
</tbody>
</table>

Source: By Author

**Figure 15. Summary statistics by outcome for SNPoors Naive Bayes Model**

The variables REER, CAB, GDPpc and HDDI have high means when the index is more stable where as PIR, BOP, UR, CPIH and FDGDP have lower mean values when the index is more stable. The findings agree with the analysis done by Mellios and Paget-Blanc (2006) and Chee et al (2015) and concluded that exchange rates are one of the variables used by rating agencies to measure a country’s creditworthiness. For CAB, there seem to be less difference between the means for less stable and those for more stable. CAB showed a lot of overlap indicating that there is potential of a lot of observations being misclassified. The results are consistent with findings by Bissoondoyal - Bheenick et al (2005), Mellios and Paget-Blanc (2006), Iyengar (2010), Afonso, Gomes and Rother (2011), Arefjevs, Girts Braslinš (2013), Sánchez-Monedero et al (2014), Kabadayi and Çelik (2015), Ivanovic et al (2015), Chee et al (2015), De Moor et al (2018) and Cantor and Packer (1996) that GDP growth is a crucial variable in determining sovereign credit ratings.

The confusion matrix for the train and test data are shown in Figure 16.
Table 16. Confusion matrix for SNPoors

<table>
<thead>
<tr>
<th>Observed</th>
<th>Train Data</th>
<th>Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level of stability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>More stable</td>
<td>Less stable</td>
</tr>
<tr>
<td>More stable</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>Less stable</td>
<td>4</td>
<td>31</td>
</tr>
<tr>
<td>Percentage misclassified</td>
<td>0.0725</td>
<td>0.0909</td>
</tr>
</tbody>
</table>

Source: By Author

Figure 16. Confusion matrix for SNPoors

For the train data the true positive for class 1 = “More stable” was 33 whereas false positive was 1, for class 0 = “Less stable”, true positives was 31 and false positive was 4 for the train data. The correctly classified observations for the train data was 92.75% whereas the percentage misclassified was 7.25%. For the test data, the percentage correctly classified was 80.91% whereas those misclassified was 9.09%. The fit was better for the train data than for the test data.

When the model with Kernel smoothing was fitted to the SNPoors data, the confusion matrix is shown in Figure 17.

Table 17. Confusion matrix for SNPoors Model with Kernel smoothing

<table>
<thead>
<tr>
<th>Observed</th>
<th>Train Data</th>
<th>Test Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level of stability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>More stable</td>
<td>Less stable</td>
</tr>
<tr>
<td>More stable</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td>Less stable</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>Percentage misclassified</td>
<td>0.0145</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Source: By Author

Figure 17. Confusion matrix for SNPoors Model with Kernel smoothing

For the train data the true positive for class 1 = “More stable” was 36 whereas false positive was 1, for class 0 = “Less stable”, true positives was 32 and false positive was 0 for the train data. The correctly classified observations for the train data were 98.55% whereas the percentage misclassified was 1.45%. For the test data, the percentage correctly classified was 100% whereas those misclassified was 0%. The fit was very good for the test data than for the train data. The error rate using SNPoors data was less than 2% which is quite good and thus, the SNPoors classification was quite high.

4.5. Summary of findings

The Machine learning tool Naïve Bayes models can classify economic variables used in rating sovereigns with precision. Naïve Bayes managed to classify the given variables using the two SCR categories that is more stable and less stable. More stable ratings signify improving, high or upgrading ratings and less stable signifies the opposite. Variables classified under more stable indicates that ratings are high or favorable and those for less stable show unfavorable or low ratings. The model highlighted variables that influence ratings under more stable scenario and those that influence less stable ratings or downgrades. For Fitch variables that induce more stable ratings were REER, GDPpc, HDDI and CPIH whereas PIR, BOP, CAB, UR and FDGDP were inducing less stable ratings. Moodys’ more stable ratings were supported by REER, GDPpc and HDDI and the less stable ratings were induced by PIR, BOP, CAB, UR, CPIH, and FDGDP. The study found that Credit Rating Agencies (CRA) use different macroeconomic variables to assess and assign sovereign ratings. The only variable that had problems of overlapping was CPIH for both Fitch and Moodys ratings. S&P ratings had similar classification as the Moodys but the only difference was that CAB was the overlapping variable.

5. Conclusion

The Naïve Bayes Classification Machine Learning model captured the relationship between SCR and economic indicators by clearly identifying crucial determinants of SCR. The findings imply that for sovereigns to avoid rating downgrades they should control household debt to favorable levels, reduce inflation, mitigate exchange rate risks and continuously maintain growth in GDP. The results confirm the findings by Bennell, Crabbe Thomas and Ap Gwilym, (2006); Kumar and Haynes, (2003); and
Kräussl, (2005) and Cantor and Packer (1996), that sovereign ratings effectively summarize and supplement the information contained in macroeconomic indicators and are therefore strongly correlated with market-determined credit spreads. Lack of data on some variables like corruption, governance and so on affected the depth of the analysis. In Africa, data is not readily available to conduct a feasible research and productive analysis. Bank data are highly concealed, confidential and private to protect the financial sector hence not accessible to researchers. Data is focused on one specific nation, South Africa which might have unique characteristics or conditions than other countries. Machine Learning models have weaknesses of misclassification and overfitting which might not precisely be corrected by cross validation. Further studies must incorporate non-quantitative variables like governance, corruption, political stability and effective management of government institutions. New models should be applied that are not affected by weaknesses observed in machine learning like overfitting data and misclassification.

References


Arefjevs, I. and Brasliņš, G. 2013. Determinants of sovereign credit ratings – example of Latvia. *New Challenges of Economic and Business Development* C41, H30


