INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002
Received: 17.10.2021 | **Accepted:** 07.01.2022 | **Published Online:** 08.03.2022

-RESEARCH ARTICLE-

AN ANALYSIS OF SOVEREIGN CREDIT RATINGS USING RANDOM FOREST

Oliver Takawira

Lecturer: Department of Finance and Investment Management (DFIM), College of Business and Economics (CBE) - University of Johannesburg, South Africa
Email address: otakawira@uj.ac.za
https://orcid.org/0000-0001-7515-1733

John W. Muteba Mwamba

Lecturer and Professor: DOE - School of Economics, College of Business and Economics (CBE) – University of Johannesburg, South Africa
Email address: johnmu@uj.ac.za; mutebamj@mcmaster.ca; moi175@hotmail.com
https://orcid.org/0000-0002-8002-1156

—Abstract—

The study's objective was to build a novel approach for analysing sovereign credit ratings. Quarterly data from 1999 to 2020 were analysed using a random forest model generated from decision trees. Macroeconomic indicators and sovereign credit ratings (SCR) ''were used from the three major credit rating agencies, Fitch, Moody's, and Standard & Poor's. The random forest model is a machine learning methodology that analyses and forecasts data using categorisation algorithms. The random forest classifier and analyser fared admirably well when classifying and analysing sovereign credit ratings. The data imply that the most relevant variables for estimating and ranking credit ratings are household debt to disposable income, exchange rates, and inflation. The data indicate that increases in economic metrics such as Real Effective Exchange Rates, Gross Domestic Product Growth, Household Debt to Disposable

Citation (APA): Takawira, O., Mwamba, J. W. M. (2022). An Analysis of Sovereign Credit Ratings Using Random Forest. *International Journal of Economics and Finance Studies, 14* (1), 29-87. doi: 10.34109/ijefs. 202220002

Income, and Consumer Price Index Headline result in ratings shift in favour of the borrower. Authorities should maintain a low HDDI, stabilised inflation, and a stronger currency to encourage sovereign rating upgrades.s

Key Words: Sovereign Credit rating, Random Forest, Decision Tree, Machine Learning and Macroeconomic variables

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

1. INTRODUCTION

Sovereign credit ratings (SCR) convey a borrowing nation's default risk. Forecasting sovereign ratings enable countries to ascertain the prospects for their default risk as regarded by lenders (B. Kabadayı, and Ahmet Alkan Çelik, 2015). Polito (2013), as well as System (2011a), noted that some scholars had accused Credit Rating Agencies (CRA) of escalating the debt crisis, increasing borrowing prices, and causing or worsening financial instability.

Developing countries must forecast sovereign credit ratings to minimise financial instability and other negative consequences of sovereign credit rating downgrades (Kiff et al., 2010). As a developing country with erratic political stability, South Africa is particularly sensitive to the negative consequences of sovereign rating downgrades (O. Takawira et al., 2021). Negative expectations or speculative actions aimed at downgrading can negatively affect macroeconomic indicators, resulting in financial instability and disruptions to the fiscal or monetary framework. CRAs lack the necessary regulation, oversight, and transparency to ensure that their services and contributions to the financial sector and economy are efficient, effective, and highly productive. Underdeveloped countries lack sovereign rating agencies to prevent being exploited by the uncontrolled and oligopolistic credit rating market (Marandola et al., 2017).

Moraes (2011) has shown that credit ratings face difficulties due to rating agencies' opaque procedures and inability to forecast financial catastrophes such as the late 1990 Asian crisis. Moraes (2011) stated that CRAs do not look ahead when assigning ratings but rather backwards. According to Ozturk et al. (2016), the 2008 financial crisis resulted from over-reliance on sovereign ratings of credit rating agencies. Ratings have a legitimate economic purpose, and local enterprises are often rated lower than the county's total sovereign rating (Iyengar, 2010).

Sovereign credit ratings affect asset interest rates and serve as the baseline for credit risk assessment for other assets, affecting the breadth and volume of assets (Ozturk et al., 2016). Ozturk et al. (2016) noted that disagreements over the veracity of sovereign ratings exacerbated the necessity for additional internal credit scoring systems to avoid an over-reliance on CRAs. Governments are keen on credit ratings to increase access to

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

international capital markets and lower borrowing costs. Sovereign ratings are not exclusively focused on governments but also on the country's other debtors (Iyengar, 2010).

Previously published research converted ratings to numerical values, obliterating certain information, but this study analysed ratings in their original category format as released by CRAs. The absence of variables such as a political stability index, governance measurement, corruption indices, and gold prices in previous research (Afonso et al., 2011; Cantor et al., 1996; G. Ferri et al., 1999; B. Kabadayı et al., 2015; Kräussl, 2005; Kumar et al., 2003; Ozturk et al., 2016) has been noted. Studies such as Bissoondoyal-Bheenick et al. (2006) and Balikçioğlu et al. (2019) that used probit regression models have the limitation that the model required normal distributions for all unobserved utility components. Logit regression studies such as Oskonbaeva (2020) and Afonso et al. (2009) were constrained by the assumption of linearity between the dependent and independent variables. This study used a random forest classification technique to analyse SCR in their native symbol format, rather than to transform them to numerical values as prior research has done (Bennell et al., 2006; Kräussl, 2005; Kumar et al., 2003; O. a. M. Takawira, W.M, 2020). Random forest and other machine learning models are unaffected by the distribution of data and unobserved components and are not constrained by the relationship between dependent or independent variables, making them unique (Briard et al., 2020). The project sought to answer the following question: Can a machine learning technique such as random forest be used to detect trends in sovereign ratings, classify them, analyse them, and forecast future rating movements?

The purpose of this study was to use a machine learning model to analyse and forecast SCRs to assist governments in avoiding downgrades and promoting upgrades that would restore financial stability (Bennell et al., 2006; Cantor et al., 1996; Kräussl, 2005; Kumar et al., 2003; O. Takawira et al., 2021). Additionally, the study evaluated the relationship between sovereign credit ratings and macroeconomic data to unearth the determinants of sovereign credit ratings from Fitch, Moody's, and Standard & Poor's (S&P). This work aims to develop a model capable of forecasting future sovereign ratings, using machine learning to classify ratings and uncovering macroeconomic aspects that CRAs consider when assessing sovereigns (O. a. M. Takawira, W.M, 2020). It is for identifying the macroeconomic elements that CRAs consider when rating sovereigns would aid policymakers and authorities in developing measures that strengthen these macroeconomic parameters and promote rating upgrades (Duke et al., 2020).

The three largest credit rating agencies (CRAs) are Standard & Poor's (S&P), Moody's, and Fitch Rating Agency. Ratings from CRAs have been rejected due to the oligopolistic nature of the rating industry. Moody's, Standard & Poor's, and Fitch all use similar, but not identical, terminology when rating bonds. They are alphabetical, with '‘AAA' being the highest S&P grade. The words' investment grade' and speculative grade' have been

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

coined to refer to the categories ranging from 'AAA' to 'BBB' (investment grade) and from 'BB' to 'D' (speculative grade) (speculative grade). The following table summarises the terminology used by rating agencies.

1.1.1.1.1.1 Table 1: S&P, Moody's and Fitch Rating Systems

	S&P	Moody's	Fitch
Prime	AAA	Aaa	AAA
High grade	AA+	Aa1	AA+
	AA	Aa2	AA
	AA-	Aa3	AA-
Upper medium grade	A+	A1	A+
	A	A2	A
	A-	A3	A-
Lower medium grade	BBB+	Baa1	BBB+
	BBB	Baa2	BBB
	BBB-	Baa3	BBB-
Non-investment grade speculative	BB+	Ba1	BB+
	BB	Ba2	BB
	BB-	Ba3	BB-
Highly speculative	B+	B1	B+
	B	B2	B
	B-	B3	B-
Substantial risks Extremely speculative - In default with little prospect for recovery	CCC+	Caa1	CCC+
	CCC	Caa2	CCC
	CCC-	Caa3	CCC-
	CC	Ca	CC; C
In Default	SD; D	C	DDD; DD; D

Chee et al. (2015); Afonso et al. (2011), Mellios et al. (2006), Hill et al. (2010).

According to Mahomed Karodia et al. (2014), "South Africa's foreign debt rating has been consistently reduced by international rating agencies over the previous several months, and downward growth revisions have become a persistent pattern." As of 2014, the country had been downgraded by all rating agencies, incurring damaging downgrades (O. Takawira, and Motseta, S.D. , 2021). As an emerging economy, South Africa is highly susceptible to loan volatility due to the country's recent downgrades, which have harmed growth and financial sector stability (Mahomed Karodia et al., 2014).

Kume (2012) claimed, using Partnoy (2003:52), that the relevance of rating information is limited because there is a strong association between ratings and actual defaults due

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

to CRAs downgrading ratings in response to new data being publicised or published. Cantor et al. (1996), G. Ferri et al. (1999), Afonso et al. (2011), and Ozturk et al. (2016) claim that macroeconomic variables drive sovereign ratings. However, Kume (2012) argues that significant macroeconomic variables do not cause credit rating changes (Gudalov et al., 2020).

The purpose of this work is to model, predict, and classify sovereign credit ratings and discover their determinants using a machine learning model called random forest. It is used to determine whether machine learning can organise, analyse, and forecast sovereign credit ratings.

2. LITERATURE REVIEW

Historically, the majority of research has analysed SCRs transformed to numerical format. Archer et al. (2007), Butler et al. (2006), Cantor et al. (1996), G. Ferri et al. (1999), Mora (2006), and Ratha et al. (2011). If a country's foreign currency debt is downgraded to 'junk status,' the cost of borrowing money in global markets increases (de Villiers et al., 2020; Meyer et al., 2021; Mutize, 2021; Slabbert, 2019; Weyers et al., 2017). If a credit downgrade results in ongoing currency weakness and inflationary pressure, prompting the central bank to raise the repo rate, house loan repayments will increase as well (Meyer et al., 2021; Mutize, 2021).

2.1 Theoretical Literature on Determinants of Sovereign Credit Rating

According to Sibanda (2018), the existing theoretical framework for sovereign credit ratings is primarily motivated by sovereign debt and default, as sovereign credit ratings consider two dimensions of debt settlement: ability and desire. Ability is the debt servicing capacity approach; the unforeseen worsening of the sovereign's ability to service its debt may eventually result in default (Habanabakize, 2020). The sustainability of debt affects the likelihood of default, and sustainability can be influenced by macroeconomic factors, economic policy, currency crises, short-term fiscal mismanagement, and internal or external shocks (Schumacher et al., 2021; Sibanda, 2018).

Sibanda (2018) and Schumacher et al. (2021) explained that readiness to pay a loan is primarily motivated by the need to maintain a positive reputation and ensure future access to international financial markets. Countries avoid defaulting on financial obligations, commitments, restructuring, or restructuring to avoid consequences in the inability to access international financing or loans (Kikulwe et al., 2020). According to economic theory, economic variables represent a country's willingness and ability to pay or service its debt. These variables include the economic growth and development rate, the overall increase in the price level of goods and services inside the borrowing country,

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

the real exchange rate, the foreign debt-to-GDP ratio, and the borrowing country's default history (Schumacher et al., 2021; Sibanda, 2018).

2.2 Empirical Literature Review

Due to the opaque nature of the methodology employed by credit rating agencies in judging creditworthiness, additional research has been conducted on credit rating. The disparities in the credit ratings assigned by CRA to comparable enterprises and sovereigns have prompted doubts about which factors genuinely determine credit ratings. Ozturk et al. (2016), citing Cantor et al. (1996) and G. Ferri et al. (1999), highlighted that early literature examined the extent to which variation in sovereign ratings could be explained by country-specific variables and identified certain macroeconomic, financial, and political variables that appear to be capable of explaining sovereign credit rating variation. In their evaluation of the literature, B. Kabadayı et al. (2015) noted that Archer et al. (2007), Butler et al. (2006), Cantor et al. (1996), G. Ferri et al. (1999), Mora (2006), and Ratha et al. (2011) all used sovereign ratings as quantitative dependent variables. Giovanni Ferri et al. (2001), Gaillard (2006), Hill et al. (2010), Gupta et al. (2017), and Mora (2006) all employed sovereign evaluations as dependent qualitative variables.

Kumar et al. (2003) compared the discriminant analysis model and Artificial Neural Networks (ANN), concluding that the latter model was superior to the former. The ANN model improves efficiency and speed in practical applications of the rating process, and with sufficient input data, the ANN model may produce an automatic rating with a high degree of reliability (Kumar et al., 2003). Hall et al. (2009) used the ANN to estimate bank credit risk in an Indonesian case study. The research forecasted the default rate of Indonesian Islamic banks using critical macroeconomic variables, including GDP growth, inflation, stock prices, exchange rates, and money in circulation. They found that stock prices may be utilised to forecast future issues. Pacelli (2011) used a technique based on artificial neural networks to manage credit risk. They concluded that traditional approaches and neural networks each have distinct advantages and disadvantages that must be carefully considered (Matthews et al., 2020). This study is unusual in that it uses random forest, a new machine learning model, to classify and forecast sovereign ratings that were previously classified and forecast using established statistical models such as probit, logit, and panel regression (Valencia, 2020).

3. RESEARCH METHODOLOGY

The study examined if macroeconomic variables in the country can explain variation in sovereign credit ratings. From 1999 through 2020, South African quarterly statistics on macroeconomic indicators and sovereign ratings were compiled from Quantec accessible data, trade economics, the SARB database, and Thomson Reuters. There were

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

84 observations divided 80:20 between the train and test sets. Macroeconomic indicators such as Real Effective Exchange Rates (REER), Gross Domestic Product (GDPpc), Household Debt to Disposable Income Ratio (HDDIR), Consumer Price Index Headline (CPIH), Prime Interest Rate (PIR), Unemployment Rate (UR), Foreign Debt to Gross Domestic Product (FDGDP), Balance of Payments (BOP), and Current Account Balance (CAB) were incorporated into the model to classify and forecast future sovereign credit ratings (SCR). Positive GDPpc, surplus BOP, and surplus CAB are favourable when FDGDP, PIR, HDDI, UR, deficit BOP, and deficit CAB are declining (O. Takawira et al., 2021; O. a. M. Takawira, W.M, 2020). The following equation demonstrates how the model formulations are constructed:

$$Sovereign\ Rating_{jit} = f(Macroeconomic\ Variables_t) + error_t \qquad [1]$$

where 'j' denotes the credit rating agency chosen from Fitch, Moody, or S&P, 't' denotes time in quarterly periods from 1999 to 2020, and I denote a dummy for a more stable rating' or a 'less stable rating'. This study concentrated exclusively on South Africa.

Cross-validation was used in this study since all variables in the data were split 80:20 across the train and test sets. 80% of the data was utilised for training the random forest models, and the remaining 20% of the test set's explanatory factors were used to estimate 20% of the test set's response variable.

3.1 Decision Trees (DT) / Classification and Regression Trees (CART) Model

To make predictions, the random forest uses a decision tree model dubbed categorisation and regression trees (CART). The decision tree model employs a technique called recursive partitioning, which divides data into many sub-classes so that outcomes are homogeneous among the sub-spaces. Under machine learning, the decision tree algorithm uses continuous variables for regression decision trees, but when the variables are categorical, the process switches to a classification decision tree.

According to Gupta et al. (2017), CART, which Breiman invented in 1984, is a method that constructs both classification and regression trees. The Gini Index is CART's splitting feature for binary splitting in classification trees (Wallenius et al., 2020). The regression tree performs regression analysis and can forecast a dependent variable over a specific period given a set of predictor factors (Gupta et al., 2017). The classification tree is used when the expected outcome is a class to which the data belongs; when the predicted outcome is an actual number or working with continuous variables, the regression tree is employed (Mohamed, 2017).

According to Mohamed (2017) and Rudd (2017), the decision tree-CART model is self-explanatory, simple to understand and interpret, requires minimal data preparation, handles both categorical and numerical data, handles multi-output problems, performs variable selection, establishes interactions between variables, handles missing values on data, handles outliers, and performs well even when dataset assumptions are violated slightly. Decision trees are aesthetically analogous to human decision-making and can easily be applied to industry regulation credit reasoning (Rudd, 2017).

An example of a decision tree structure is shown below in Figure 1.

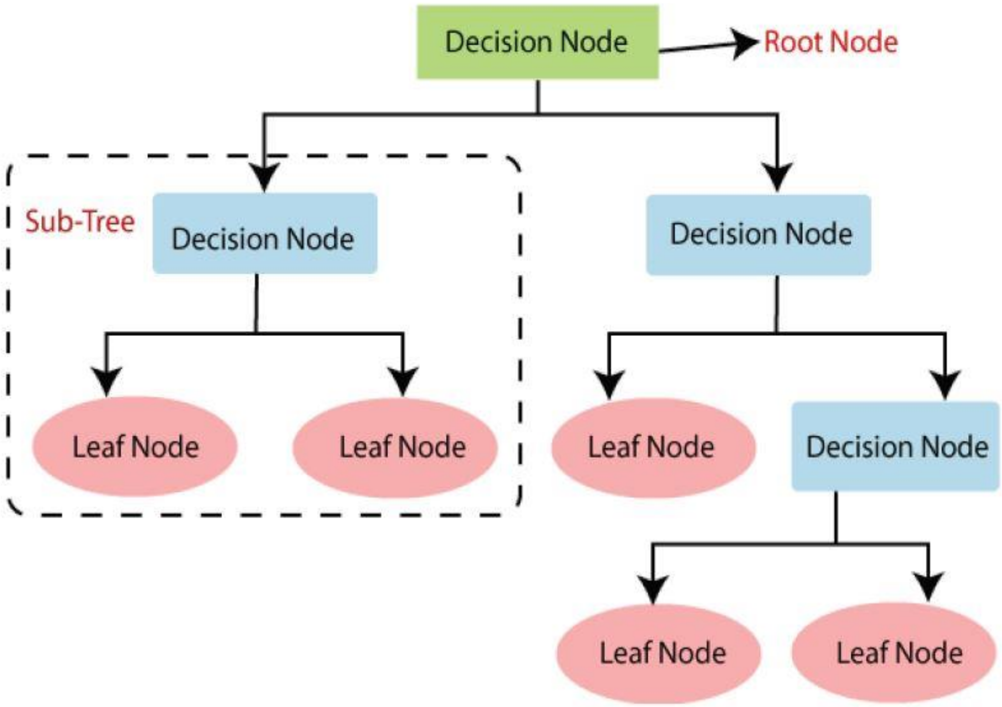


Figure 1: Example of a structure of decision tree
Source: Charbuty et al. (2021)

A decision tree is used to solve a bivariate classification problem by posing a question requiring a Yes/No response and then breaking the trees into subtrees. The initial point is a root node, and it is divided into decision nodes, which are considered parent nodes and are further divided into sub-nodes (Yoon et al., 2020). The terminal node is the leaf node. Thus, decision trees classify nodes ascending from root to leaf/terminal node. There are two methods for constructing a decision tree: Univariate and Multivariate Decision Trees (Charbuty et al., 2021).

3.2 Univariate Decision Tree

Splitting is accomplished with this technique by utilising a single characteristic on internal nodes (Mathuria, 2013). The following are some fundamental steps for constructing a tree:

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

- Determine whether all cases belong to the same class; if so, the tree is a leaf labelled with that class.

- Calculate the information and information gained for each attribute.

3.3 Determine the Optimal Splitting Attribute (Depending Upon Current Selection Criterion) Counting Information Gain

The process of calculating information gain makes use of "Entropy," a measure of data disorder, dataset impurity, or unpredictability (Bellotti et al., 2011; Charbuty et al., 2021; Hastie, 2008; Mathuria, 2013; Steinberg, 1998; Zacharis, 2018). Bits, nats, or bans are used to quantify entropy. This is sometimes referred to as quantifying the uncertainty associated with any random variable. Assume there is a fair coin; its entropy will be one bit if it is tossed once. Two bits of entropy will be generated by a series of two fair coin tosses. Entropy is always between 0 and 1. Its value is optimal when equal to 0, and is detrimental when equal to 1, i.e., the closer its value is to 0, the better (Bellotti et al., 2011; Charbuty et al., 2021; Hastie, 2008; Mathuria, 2013; Steinberg, 1998; Zacharis, 2018). The quantity of information that should be expected when viewing the output of a random variable X:

$$H(X) = E\big(I(X)\big) = \sum_i p(x_i)I(x_i) = -\sum_i p(x_i)log_2 p(x_i) \qquad [2]$$

Conditional information entropy has an equation:

$$H(X|Y) = -\sum_j p(y_j)H\big(X|Y = y_j\big)$$

$$= -\sum_j p(y_j)\sum_i p(x_i \mid y_j)log_2 p(x_i \mid y_j) \qquad [3]$$

3.4 Information Gain (IG)

If the base is 2, the unit of entropy measurement is bits; if the base is 10, the unit is dits. The term "Information Gain" refers to the relationship between inputs and outputs. It is a state-to-state shift in the entropy of information. One criterion for determining which property (feature) should be the first node is information gain (Charbuty et al., 2021; Gupta et al., 2017; Mathuria, 2013; Zacharis, 2018). The next branch has fewer qualities (features), and so on. The gain in information is calculated by subtracting the information after the split from the data before to the split (input before the division) – (information after split). Finally, the gain in data can be estimated as follows:

$$\text{IG}\ (X, Y) = H(X) - H(X|Y) \qquad [4]$$

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

3.5 Gini Index

According to Tangirala (2020), the Gini index determines the purity of a specific class after splitting along a particular attribute. The best split increases the purity of the sets resulting from the split. At node t the Gini index is defined:

$$Gini(t) = 1 - \sum_{j=0}^{1} \left(\frac{n(j|t)}{n(t)} \right)^2 \qquad [5]$$

where j is a class of target variable (in this study $j = 0$ means failure and $j = 1$ denotes success), $n(j|t)$ is the number of records of node t belonging to class j, and $n(t)$ is the total record number in node t. When the data in a node are equally distributed between all classes, the Gini index attains its maximum impurity value .5. If all data belong to the same category, the node has minimum impurity, and the Gini index is 0. To decide which attribute to split upon, the tree growing algorithm calculates the weighted average of the Gini index for the descended nodes:

$$Gini(t)_{split} = \frac{n(t_L)}{n(t)} Gini(t_L) + \frac{n(t_R)}{n(t)} Gini(t_R) \qquad [6]$$

Where t_L and t_R are the left and right child nodes of node t_R . The attribute that minimises the $Gini(t)_{split}$ is chosen to split the node.

3.6 Pruning

Pruning is a critical method to employ during tree establishment due to outliers. Additionally, it addresses overfitting. Datasets may contain sparsely specified subsets of instances. Multivariate DT has a high degree of generalisation when dealing with attribute correlation, and its output is also understandable to humans. When working with Univariate DTs, they test a single attribute several times, which in some cases results in an inefficient tree. Multivariate DT tests the data by incorporating several details into test leaves.

3.7 Random Forest (RF)

The random forest (RF), alternatively referred to as the ensemble decision tree approach, is a supervised learning classification model that constructs a forest from a collection of trees. It is suitable for classification and regression (Rai, 2017). More robustness is obtained by having a large number of bootstrapped trees in the forest, and so the Random Forest classifier produces highly accurate results when the forest has a large number of trees.

A Random Forest is a collection of decision trees assembled randomly. Random forest algorithms synthesise numerous trees by combining many different decision trees. The coupled trees are trained for random feature selection using an internal bootstrapping

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

mechanism that does not enable interaction between decision trees. It detects outliers and anomalies in expert data. According to Mohamed (2017), RF is one of the most accurate learning algorithms known, producing accurate, highly classifiers for many datasets.

Random Forest makes use of bagging, an abbreviation for "bootstrap aggregation," a technique for lowering the variance of an estimated prediction function (Kulkarni et al., 2016). The concept is to employ numerous versions of a predictor or classifier to choose a plurality vote among the predictors. In bagging, it has been demonstrated that as the number of predictors increases, accuracy increases as well until it reaches a point when it begins to decline. The highest accuracy will be achieved by generating the ideal number of predictors. Random Forests are created by combining bagging and ID3 concepts.

Each tree in the forest is generated as follows: given a training set, a random subset is sampled (with replacement) and utilised to construct a tree that resembles the ID3 concept. The Out-of-Bag (OOB) data are used to calculate the classification error rate when new trees are introduced to the forest and quantify the input variable's relevance (feature). After completing the forest, a case can be classified by a majority vote of all trees in the forest, similar to the bootstrap aggregating concept. Random Forest mixes two or more trees in unique ways, and so the process of one tree is combined with the operations of the other trees to increase the robustness of the answer.

3.8 Random Forest Algorithm

To develop the Random Forest algorithm according to Hastie and Tibshirani (2008), take $b = 1 \; to \; B$ and draw a bootstrap sample Z^* of size N from the training data, grow a random-forest tree T_b to the bootstrapped data by re-cursively repeating the following steps for each terminal node of the tree until the minimum node size n_{min} is reached, select m variables at random from the p variables, pick the best variable/split-point among the m, split the node into two daughter nodes, and the output the ensemble of trees $\{T_b\}_1^B$

To predict the new point x:

Regression: $\hat{f}_{rf}^{B}\,(x) = \frac{1}{B}\sum_{b=1}^{B} T_b\,(x)$ [7]

Classification: Let $\hat{C}_b(x)$ be the class prediction of the bth random-forest tree.

Then $\hat{C}_{rf}^{B}(x)$ = majority vote $\{\hat{C}_b(x)\}_1^B$

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

An average of B i.i.d. random variables, each with variance σ^2, has variance $\frac{1}{B}\,\sigma^2$ if the variables are simply i.d. (identically distributed, but not necessarily independent) with positive pairwise correlation ρ, the variance of the average is

$$\rho\sigma^2 + \frac{1-\rho}{B}\,\sigma^2 \quad [8]$$

As B increases, the second term disappears, but the first remains, and hence the size of the correlation of pairs of bagged trees limits the benefits of averaging. The idea in random forests is to improve the variance reduction of bagging by reducing the correlation between the trees without increasing the variance too much. This is achieved in the tree-growing process through random selection of the input variables.

Typically values for m are $\sqrt{p}$ or even as low as 1. After B, such trees $\{T(x;\ \Theta_b)\}_1^B$ are grown, the random forest (regression) predictor is:

$$\hat{f}_{rf}^{B}\,(x) = \frac{1}{B}\sum_{b=1}^{B} T(x;\ \Theta_b) \quad [9]$$

Where: Θ_b characterises the bth random forest tree in split variables, cutpoints at each node, and terminal-node values. Intuitively, reducing m will reduce the correlation between any pair of trees in the ensemble, hence reducing the variance of the average.

3.9 Empirical Analysis

Random Forest was designed to classify and analyse sovereign credit ratings from the three major credit rating agencies, Fitch, Moody's, and Standard & Poor's, utilising outlooks, micro and macroeconomic variables. REER, PIR, BOP, CAB, UR, GDPpc, HDDI, FDGDP, and CPIH were employed to forecast micro and macroeconomic indicators. Classification aids in determining which of these macroeconomic indicators are weighted most heavily in sovereign rating calculations. The classification was divided into two categories based on the credit rating agency's credit ratings: less stable and more stable. In South Africa, most ratings are upper-medium grade, lower-medium grade, or speculative non-investment grade.

The credit rating agency determines groupings into more diminutive and more stable categories. Fitch's credit rating data was classified into less stable (BBp, BB, and BBBn) categories with 30 observations and more stable categories with 50 observations (BBB and BBBp). Moody's credit rating agency assessed the Baa2 and Baa3 ratings as less stable, with 41 observations, and the Baa1 and A3 ratings as more stable, with 39 observations. The SNPoor credit rating data was segmented into less stable (BBp and BBBn) and more stable (BBB and BBBp) categories, comprising 36 and 44 observations. The Random Forest results for Fitch, Moody's, and SNPoor are reported in the following subsections.

3.10 Findings

The first model building process yielded the numbers presented in Appendix Section A for the number of trees and variables randomly picked as candidates at each split (mtry). When considering the effect of size on error, it is worth noting that for Fitch, the error rate remained constant until 300 trees, and the optimal parameter for mtry was 3. A tuned random forest model with two parameters set to 300 and 3 was fitted to the data in this scenario. Moody's error rate remained constant at approximately 500 trees (the default), and the optimal mtry parameter was 3. A tuned random forest model was fitted to the data with 500 trees and mtry set to 3, whereas for SNPoor, the effect size on the error was constant from 400, and the optimal parameter for mtry was set at 3. A tuned random forest model with ntree 400 and mtry 3 fitted the data. The fitted models yielded the following out-of-bag confusion matrix values, summarised in Table 2.

Table 2: Out-Of-Bag Error Rate and Confusion Matrix Results for Train Data and Test Data

	Train Data		**Test Data**		**Out of bag confusion matrix**
	Level of stability		**Level of stability**		
Observed	***More stable***	***Less stable***	***More stable***	***Less stable***	
Fitch More stable	41	2	7	0	5.8%
Less stable	2	24	0	4	
Percentage misclassified	4.65%	7.69%	0%	0%	
Moody's					
More stable	27	1	11	0	4.92%
Less stable	1	31	0	8	
Percentage misclassified	3.57%	6.6%	0%	0%	
SNPoor					
More stable	36	1	6	0	2.9%
Less stable	1	31	1	4	
Percentage misclassified	2.7%	3.13%	0%	20%	

Source: By Author

For the Fitch train data, the error associated with predicting less stable values was more significant than that associated with predicting more stable values. The predictions are pretty accurate, with an accuracy of 94.35% for more stable predictions and 92.31% for less stable predictions. The out of bag error rate was 5.8%, resulting in a 94.2% accuracy; this is the error rate determined using data not used in the model. The overall statistics accuracy and the Kappa coefficient for the model's data were 100% for the train and test data.

While the inaccuracy associated with forecasting less stable was more significant than that associated with expecting more durable, the forecasts are relatively accurate, with the accuracy for predicting more stable being 96.43% and that for predicting less stable being 93.94%. The out of bag error rate was 4.92%, resulting in a 95.08% accuracy. The overall statistics accuracy and the Kappa coefficient for the model's data were 100% for the train and test data.

For the SNPoor train data, the error associated with forecasting less regular trains was more significant than the error associated with predicting more regular trains. The predictions are pretty accurate, with an accuracy of 97.30% for more stable predictions and 96.87% for less stable predictions. The out of bag error rate was 2.9%, resulting in a perfect accuracy of 97.1%. The overall statistics accuracy and the Kappa coefficient for the train data were 100%; however, the accuracy was 19.91% for the test data, and the Kappa coefficient was 81.36%.

The metrics of predictor variable relevance are beneficial for variable selection and model interpretation. The Random Forest gives measures of varying relevance that aid in determining the most critical variables for differentiating between less stable and more stable credit ratings. The model's critical variables are listed in Tables 3, 4, and 5 for Fitch, Moody's, and SNPoor, respectively. Graphs depicting variable relevance are included in the Appendix section.

Table 3: Variable Importance for Fitch

Variable	Less Stable	More Stable	Mean Decrease Accuracy	Mean Decrease Gini
HDDI	14.0423	12.9927	14.9075	8.1313
CPIH	12.5893	10.5552	12.2814	6.2200
REER	11.1866	10.6710	12.8984	4.9856
BOP	7.2925	4.5035	7.4148	2.8110
UR	6.2230	7.0722	7.4260	2.9614
CAB	5.9074	4.0056	6.1289	1.6567
PIR	5.5089	5.1048	6.2362	1.7287
FDGDP	5.1088	5.4343	6.0478	1.9250

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

GDPpc	4.9008	3.0757	5.1629	1.3344

Source: By Author

According to the measured prediction accuracy, the most relevant factors for the Fitch data are HDDI, CPIH, REER, and BOP, which is consistent with Meyer et al. (2021) and Mellios et al. (2006). Table 4 summarises the variables that are critical to Moody's.

Table 4: Variable Importance for Moody's

Variable	Less Stable	More Stable	Mean Decrease Accuracy	Mean Decrease Gini
HDDI	19.9649	22.6695	23.5401	9.6278
CPIH	13.1714	14.4170	15.3407	5.0451
UR	8.1655	12.7674	13.2141	3.9985
FDGDP	7.3207	13.1574	12.5285	4.4858
REER	6.7748	7.4460	8.4312	2.0052
PIR	6.1573	4.9668	6.8175	0.8423
GDPpc	6.0776	8.0437	8.7334	1.6159
CAB	4.6961	5.2734	5.8600	1.2445
BOP	2.4523	3.4627	3.8609	0.8703

Source: By Author

According to Meyer et al. (2021) and Mellios et al. (2006), the essential factors for Moody's data are HDDI, CPIH UR, and FDGDP, whereas, for SNPoor, the most critical variables are shown in Table 5.

Table 5: Variable Importance for SNPoor

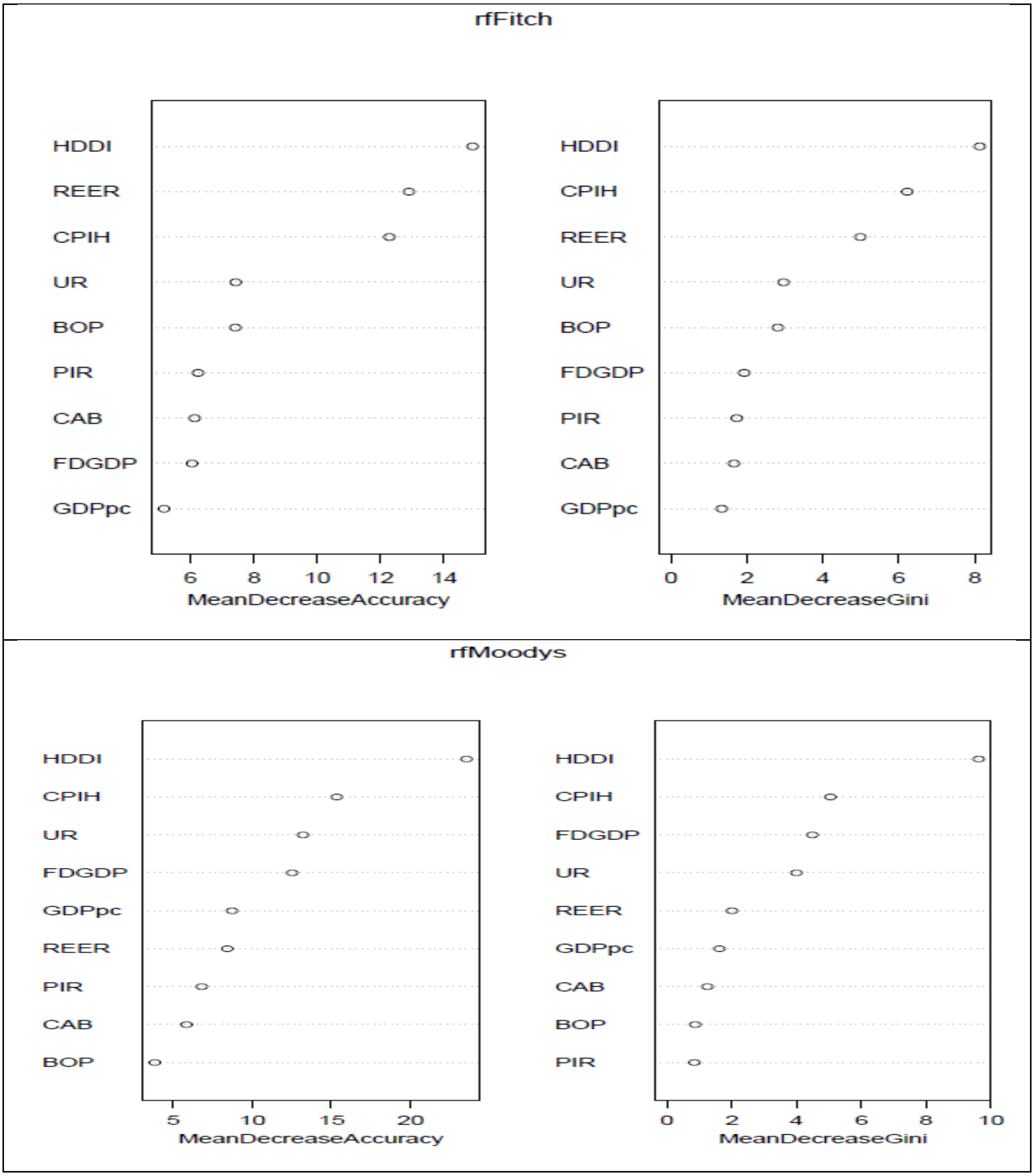
Variable	Less Stable	More Stable	Mean Decrease Accuracy	Mean Decrease Gini
HDDI	18.5805	16.4116	20.0304	8.1779
REER	17.3257	16.7349	19.2835	9.2001
CPIH	12.5587	12.7967	13.6406	5.7131
GDPpc	8.4285	7.1177	8.7814	2.8162
CAB	7.3526	2.5343	6.9501	1.5985
PIR	6.9905	5.9168	8.1316	1.5713
UR	5.4415	5.5456	7.0372	2.2189
BOP	4.7455	1.4491	4.2863	0.9568
FDGDP	3.9371	5.5482	6.0444	1.5321

Source: By Author

For the SNPoor data, the most important variables were HDDI, REER, CPIH and GDPpc, as confirmed by Meyer et al. (2021) and Mellios et al. (2006).

3.11 Variable Importance Graphs

Looking at the three data sets, the essential variables in all the data sets are HDDI and CPIH, while REER was necessary for the Fitch and SNPoor data set. The vital variables are shown in Figure 2 for the three data sets.



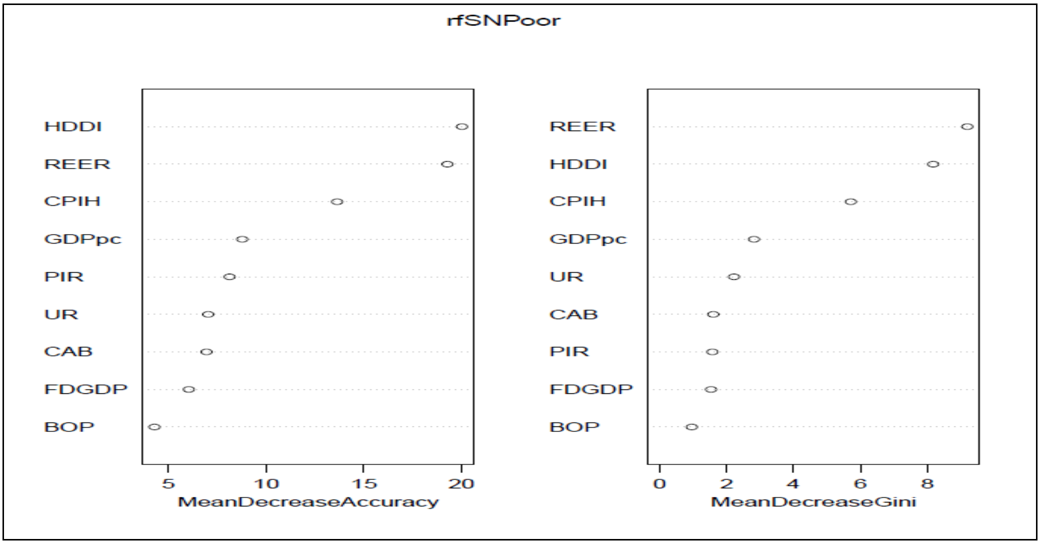


Figure 2: Variable Importance Graphs
Source: By Author

3.12 Predicting Future Sovereign Credit Ratings

The projected models for Fitch, Moody's, and SNPoor were used to test if the model accurately categorised observations not included in the study, namely quarterly data from the first through second quarters of 2020. All of the observations are listed in Table 6 below. They all fall into the "less stable" category for all credit rating agencies.

The next sections will discuss the random forest trees generated from each data source. The random forest produced 500 trees, of which the first six were used. The results part contains the initial trees and prediction tables for each credit rating agency (Fitch, Moody's, and S&P), while the Appendix section has the remaining tree diagrams and predictions for each credit rating agency.

3.13 Prediction Using Fitch Model

Figure 3 gives the tree generated by the Fitch model. Figure 3 below shows the predicted class for Fitch using tree one.

The classification of the future observations for Fitch model tree 1 was done using the variables CPIH, FDGDP, REER, and PIR using five paths.

Path 1: If CPIH is less than 50.4664, the observation is classified as less stable.

Path 2: If CPIH is greater than 50.4664 and FDGDP is less than 13.4115, the classification is more stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 28-87) **Doi:** 10.34109/ijefs. 202220002

Table 6: Quarterly Data from 2019 to Mid-2020

Time	GDPpc	FDGDP	CAB	BOP	REER	PIR	HDDI	UR	CPIH	Fitch Class	Moodys Class	SNPoors Class
2019/03/31	0	17.04585	-57063.9	42301	93.57	10.25	73	27.6	110.1	0	0	0
2019/06/30	0.9	16.04835	-29185.8	-31495	92.01333	10.25	72.9	29	112.0333	0	0	0
2019/09/30	0.1	17.48545	-62178.4	43992	91.33667	10	72.6	29.1	113.1	0	0	0
2019/12/31	-0.5	19.66327	57.3302	102486	91.28667	10	73.2	29.1	113.5667	0	0	0
2020/03/31	0.1	21.02702	-8168.36	201663	88.19333	9.416667	73.6	30.1	114.9667	0	0	0
2020/03/30	-17.1	26.17012	-8503.1	91508	76.50333	7.416667	85.3	23.3	114.7333	0	0	0

Source: By Author

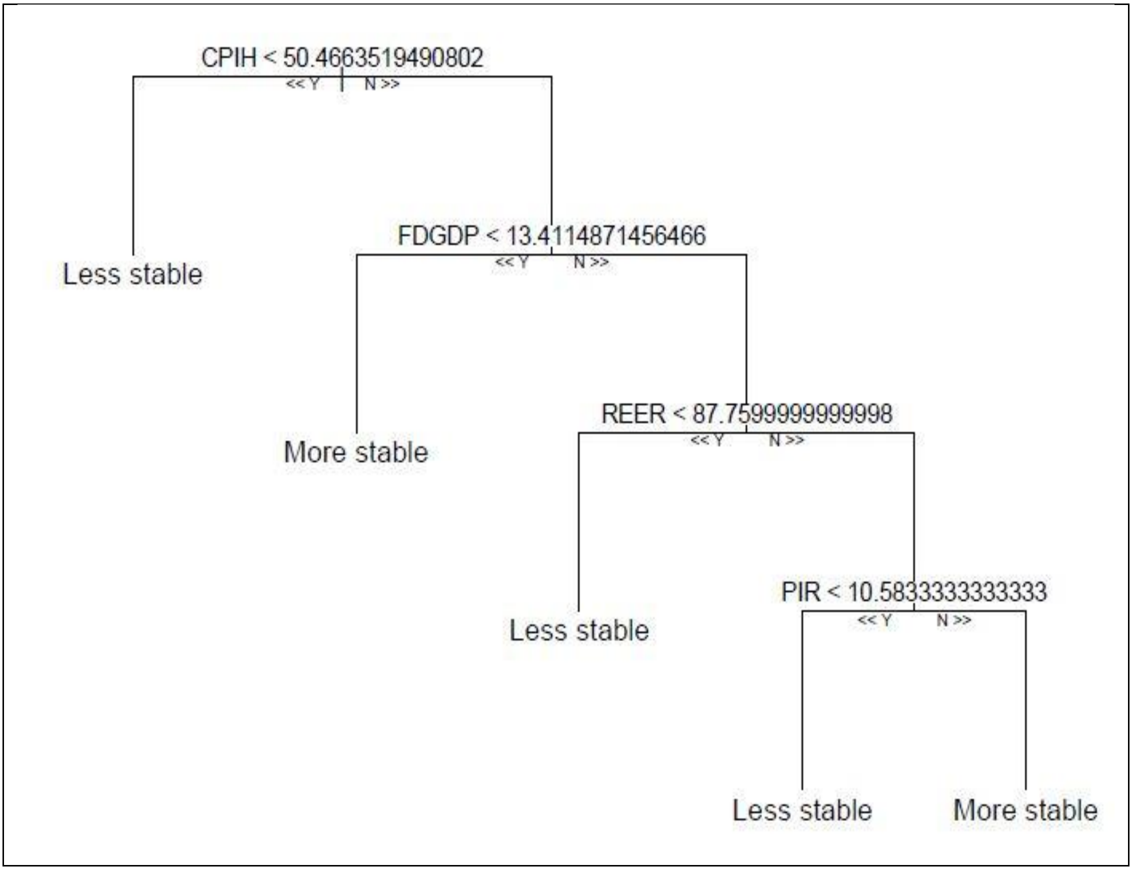


Figure 3: Prediction tree for Fitch Model Tree 1
Source: By Author

Path 3: If CPIH is greater than 50.4664, FDGDP is more than 13.4115, and REER is less than 87.76, the classification is less stable.

Path 4: If CPIH is greater than 50.4664, FDGDP is more than 13.4115, REER is more than 87.76, and PIR is less than 10.5833, then the classification is less stable.

Path 5: If CPIH is greater than 50.4664, FDGDP is more than 13.4115, REER is more than 87.76, and PIR is more than 10.5833, then the classification is more stable.

Observations 1 to 5 had CPIH greater than 50.4664, FDGDP more than 13.4118, REER more than 87.76, and PIR less than 10.5833, classified as less stable. Observation 6 had CPIH more than 50.4664, FDGDP more than 13.4118, REER less than 87.76, and was ranked as less stable. Thus, using the tree generated in Figure 3, the predicted class and the current class are shown in Table 7.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Table 7: Fitch Prediction Model for Tree 1

Time	CPIH	FDGDP	REER	PIR	Fitch Class	Predicted Fitch Class Tree 1
2019/03/31	110.1	17.04585	93.57	10.25	0	0
2019/06/30	112.03333	16.04835	92.01333	10.25	0	0
2019/09/30	113.1	17.48545	91.33667	10	0	0
2019/12/31	113.56667	19.66327	91.28667	10	0	0
2020/03/31	114.96667	21.02702	88.19333	9.416667	0	0
2020/06/30	114.73333	26.17012	76.50333	7.416667	0	0

Source: Author

The first tree model for Fitch correctly predicted the classification of all the observations; that is, the expected class was less stable (“0”).

3.14 Summary of the Prediction of Fitch Model

By examining the findings, it is clear that the model misclassified the last observation in the scenario where the variable HDDI was included in the tree diagram. As was the case in our earlier study (O. a. M. Takawira, W.M, 2020), HDDI was the most significant variable in the model when considering mean decrease accuracy and mean drop Gini. One can deduce that the first five observations are projected to be less stable than the last observation. The results indicated that five observations were correctly classified as “more stable” by all tree diagrams, but the final observation was incorrectly classified as “less stable” when the variable HDDI was included in the model. However, the prediction results indicated that the fitted Fitch model was adequate.

3.15 Prediction Using Moodys Model

Figure 4 gives the tree generated by Moody’s model. Figure 4 below shows the predicted class for Moody’s using tree one.

Moody’s model tree 1 was constructed using the variables UR and CPIH using four paths.

Path 1: If UR is less than 25.9 and CPIH is less than 48.9136, the classification is less stable.

Path 2: If UR is less than 25.9, CPIH is more than 48.9136, and CPIH is less than 89.0167, the classification is stable.

Path 3: If UR is less than 25.9, CPIH is more than 48.9136, and CPIH is more than 89.0167, the classification is less stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

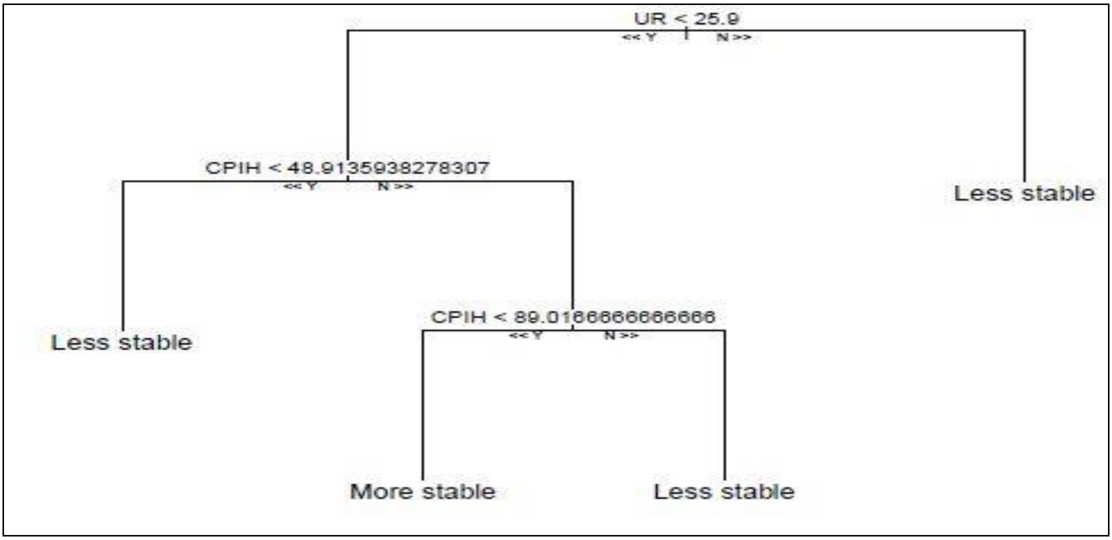


Figure 4: Prediction tree for Moodys Model Tree 1
Source: By Author

Path 4: If UR is more than 25.9, the classification is less stable.

Observations 1 to 5 had UR more than 25.9 and were classified as less stable. Observations 6 had UR less than 25.9, CPIH more than 48.9136 and CPIH more than 89.0167 and was classified as less stable. Using the tree generated in Figure 3, the predicted class and the current class are shown in Table 8.

Table 8: Moody's Prediction Model for Tree 1

Time	UR	CPIH	Moodys Class	Predicted Moodys Class Tree 1
2019/03/31	27.6	110.1	0	0
2019/06/30	29	112.0333	0	0
2019/09/30	29.1	113.1	0	0
2019/12/31	29.1	113.5667	0	0
2020/03/31	30.1	114.9667	0	0
2020/06/30	23.3	114.7333	0	0

Source: On Author

The first tree model for Moody correctly predicted the classification of all the observations; that is, the expected class was less stable ("0").

3.16 Summary of the Prediction of Moody's Model

Three of the models misclassified the final observation in some instances when the variables HDDI and PIR were included in the tree diagram. When comparing mean decrease accuracy and mean decrease Gini, HDDI was one of the more significant variables in the model than the other variables, corroborating the findings from our prior work (O. a. M. Takawira, W.M, 2020). One can deduce that the first five observations are projected to be less stable than the last observation. The results indicated that five observations were accurately classified as "more stable" in all tree diagrams, but the final observation was incorrectly classified as "less stable" in several models that included HDDI and PIR. The prediction results show that Moody's model that was fitted was adequate.

3.17 Prediction Using SNPoor Model

Figure 5 gives the tree generated by the SNPoor model. Figure 5 below shows the predicted class for SNPoor using tree one.

The SNPoor model tree 1 classified the observations using five paths through the variables, CPIH, PIR, and BOP.

Path 1: If CPIH is less than 85.1 and PIR is less than 14.1667, the observation is classified as more stable.

Path 2: If CPIH is less than 85.1, PIR is more than 14.1667, and CPIH is less than 51.0033, the observation is less stable.

Path 3: If CPIH is less than 85.1, PIR is more than 14.1667, CPIH is more than 51.0033, and BOP is less than 34119, then the observation is classified as more stable.

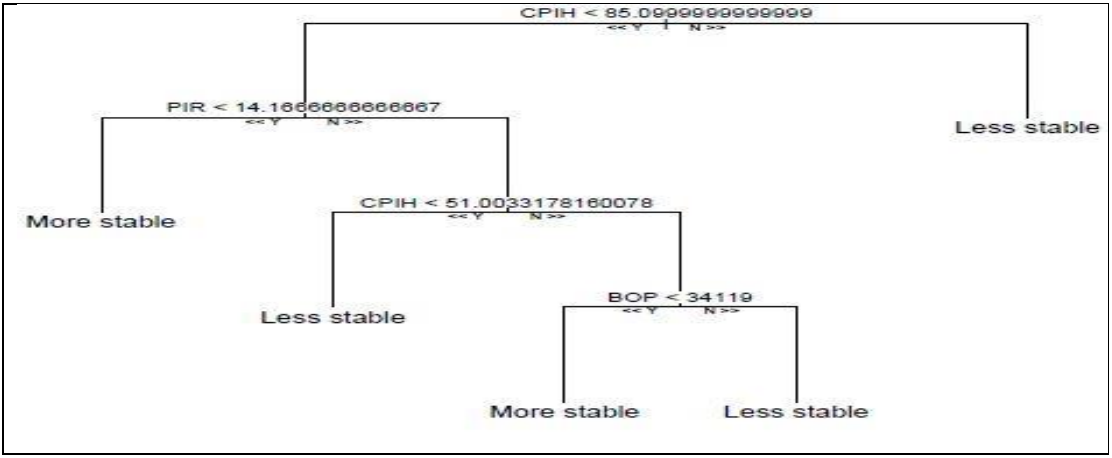


Figure 5: Prediction tree for SNPoor Model Tree 1
Source: By Author

Path 4: If CPIH is less than 85.1, PIR is more than 14.1667, CPIH is more than 51.0033, and BOP is more than 34119, the observation is classified as less stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Path 5: If CPIH is more than 85.1, the classification is less stable.

All the observations had CPIH as more than 85.1 and were classified as less stable. Using the tree generated in Figure 5, the predicted class and the current class are shown in Table 9.

Table 9: SNPoor Prediction Model for Tree 1

Time	CPIH	PIR	BOP	SNPoor Class	Predicted SNPoor Class Tree 1
2019/03/31	110.1	10.25	42301	0	0
2019/06/30	112.0333	10.25	-31495	0	0
2019/09/30	113.1	10	43992	0	0
2019/12/31	113.5667	10	102486	0	0
2020/03/31	114.9667	9.416667	201663	0	0
2020/06/30	114.7333	7.416667	91508	0	0

Source: On Author

The first tree model for SNPoor correctly predicted the classification of all the observations; that is, the expected class was less stable (“0”).

3.18 Summary of the SNPoor Model

By examining the findings, it is clear that the model misclassified the last observation in the scenario where the variable HDDI was included in the tree diagram. When mean decreases accuracy and mean decrease Gini were considered, HDDI was the more significant variable in the model, consistent with O. a. M. Takawira, W.M (2020). The third model, in which four observations were misclassified, was discarded since the deciding variable did not rank among the top four in terms of variable relevance in PIR. The remaining five models demonstrate that the first five observations are projected to be less stable than the last observation. All five tree diagrams accurately classified five observations; however, the final observation was incorrectly classified as “less stable” when the variable HDDI was included in the model. The prediction findings show that the SNPoor model that was fitted was adequate.

4. CONCLUSION

By explicitly identifying critical factors of SCR, classification under machine learning captures the relationship between SCR and economic indices. According to the study, both micro and macroeconomic variables affect sovereign ratings and are significant variables considered by credit rating agencies in rating sovereigns. In comparison to other variables, Random Forest revealed that HDDI and CPIH were the most frequently used variables in determining SCR from the three major CRAs. REER was beneficial for Fitch and SNPoors ratings, whereas UR and FDGDP were beneficial for Moody's ratings. Exchange rates are beneficial to ratings because a more robust currency enables a country to repay foreign debts more quickly than if borrowings were made in a weaker currency. Unemployment mimics a lack of government revenue; if more residents are working, they contribute to government revenue via taxes, and so CRAs may upgrade the country's sovereign rating in the belief that the increased money may be used to service government debt.

The findings show that sovereigns should keep household debt at manageable levels, contain inflation, moderate exchange rate concerns, and sustain continuous GDP growth to avoid rating downgrades. The findings corroborate those of Bennell et al. (2006); Kumar et al. (2003); and Kräussl (2005), as well as Cantor et al. (1996), that sovereign ratings effectively summarise and supplement the information contained in macroeconomic indicators and are thus highly correlated with market-determined credit spreads. Additionally, the findings corroborate those of our prior research, in which we used Naive Bayes to analyse sovereign credit ratings (Mellios et al., 2006; Meyer et al., 2021; O. Takawira et al., 2021). However, the results may vary when a large data sample is used.

Machine Learning approaches such as random forest, and decision trees can categorise, analyse, model, recognise patterns or trends, and forecast credit scores efficiently and effectively. The random forest machine learning system anticipated values with a higher degree of precision. On the other hand, credit rating agencies use distinct techniques and economic considerations when grading sovereigns. The most critical indicators that authorities, policymakers, government officials, and central banks should watch in sovereign credit ratings are HDDI, CPIH, and REER. Increased disposable incomes encourage spending, saving, and investing, all of which contribute to economic growth. As the economy grows, the country becomes more capable of repaying borrowed monies, resulting in CRAs upgrading the country's sovereign rating. Authorities should maintain a low HDDI, low stabilised inflation, and a stronger currency to encourage sovereign rating upgrades. For future research, we advocate the inclusion of characteristics like corruption and governance in the classification of sovereign ratings and the forecasting of big data sets using machine learning models such as a random forest.

REFERENCES

Afonso, A., Gomes, P., & Rother, P. (2009). Ordered response models for sovereign debt ratings. *Applied Economics Letters, 16*(8), 769-773. doi:https://doi.org/10.1080/13504850701221931

Afonso, A., Gomes, P., & Rother, P. (2011). Short- and long-run determinants of sovereign debt credit ratings. *International Journal of Finance & Economics, 16*(1), 1-15. doi:https://doi.org/10.1002/ijfe.416

Archer, C. C., Biglaiser, G., & DeRouen, K. (2007). Sovereign Bonds and the "Democratic Advantage": Does Regime Type Affect Credit Rating Agency Ratings in the Developing World? *International Organization*, *61*(2), 341-365. doi:https://doi.org/10.1017/S0020818307070129

Balikçioğlu, E., & Yilmaz, H. H. (2019). How fiscal policies affect credit rates: probit analysis of three main credit rating agencies' sovereign credit notes. *Transylvanian Review of Administrative Sciences, 56*(15), 5-22. Retrieved from http://www.transylvanianreviewjournal.com/ojsfolder/index.php/TRAS/article/view/752

Bellotti, T., Matousek, R., & Stewart, C. (2011). A note comparing support vector machines and ordered choice models' predictions of international banks' ratings. *Decision Support Systems, 51*(3), 682-687. doi:https://doi.org/10.1016/j.dss.2011.03.008

Bennell, J. A., Crabbe, D., Thomas, S., & Gwilym, O. a. (2006). Modelling sovereign credit ratings: Neural networks versus ordered probit. *Expert Systems with Applications*, *30*(3), 415-425. doi:https://doi.org/10.1016/j.eswa.2005.10.002

Bissoondoyal-Bheenick, E., Brooks, R., & Yip, A. Y. N. (2006). Determinants of sovereign ratings: A comparison of case-based reasoning and ordered probit approaches. *Global Finance Journal, 17*(1), 136-154. doi:https://doi.org/10.1016/j.gfj.2006.06.001

Briard, R., Bhuiyan, N., Sicotte, H., & Keshani, P. (2020). Critical Success Factors in New Product Development Projects in a Weak Matrix Structure: An Aerospace Case Study. *The Journal of Modern Project Management*, *8*(2). Retrieved from https://journalmodernpm.com/index.php/jmpm/article/view/JMPM02403

Butler, A. W., & Fauver, L. (2006). Institutional Environment and Sovereign Credit Ratings. *Financial Management, 35*(3), 53-79. doi:https://doi.org/10.1111/j.1755-053X.2006.tb00147.x

Cantor, R., & Packer, F. (1996). Determinants and impact of sovereign credit ratings. *Economic policy review, 2*(2), 37-54. doi:https://dx.doi.org/10.2139/ssrn.1028774

Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, *2*(01), 20-28. Retrieved from https://www.jastt.org/index.php/jasttpath/article/download/65/24

Chee, S. W., Fah, C. F., & Nassir, A. M. (2015). Macroeconomics determinants of sovereign credit ratings. *International Business Research, 8*(2), 42.

de Villiers, C., Cerbone, D., & Van Zijl, W. (2020). The South African government's response to COVID-19. *Journal of Public Budgeting, Accounting & Financial Management, 32*(5), 797-811. doi:https://doi.org/10.1108/JPBAFM-07-2020-0120

Duke, E. O., & Osim, S. E. (2020). The culture of slavery in traditional context and globalised society. *Social Space Journal. eu*, 145. Retrieved from https://www.researchgate.net/profile/Emmanuel-Duke/publication/348431950

Ferri, G., Liu, L.-G., & Majnoni, G. (2001). The role of rating agency assessments in less developed countries: Impact of the proposed Basel guidelines. *Journal of Banking & Finance, 25*(1), 115-148. doi:https://doi.org/10.1016/S0378-4266(00)00119-9

Ferri, G., Liu, L.-G., & Stiglitz, J. E. (1999). The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis. *Economic Notes, 28*(3), 335-355. doi:https://doi.org/10.1111/1468-0300.00016

Gaillard, N. (2006). Determinants of Moody's and S&P's subsovereign credit ratings. *World Bank.*

Gudalov, N. N., & Treshchenkov, E. Y. (2020). The Resilience of the EU Neighbours to the South and to the East: A Comparative Analysis. *Croatian International Relations Review, 26*(86), 6-41. doi:https://doi.org/10.37173/cirr.26.86.1

Gupta, B., Rawat, A., Jain, A., Arora, A., & Dhami, N. (2017). Analysis of various decision tree algorithms for classification in data mining. *International Journal of Computer Applications, 163*(8), 15-19.

Habanabakize, T. (2020). Assessing the impact of interest rate, catering, and fast-food income on employment in the social services industry. *International Journal of Economics and Finance, 12*(2), 534-550. doi:https://doi.org/10.34109/ijefs.202012218

Hall, M. J. B., Muljawan, D., Suprayogi, & Moorena, L. (2009). Using the artificial neural network to assess bank credit risk: a case study of Indonesia. *Applied Financial Economics, 19*(22), 1825-1846. doi:https://doi.org/10.1080/09603100903018760

Hastie, T. a. T., R.I. (2008). The elements of statistical learning: data mining, inference, and prediction/by Trevor Hastie, Robert Tibshirani, Jerome Frieman.

Hill, P., Brooks, R., & Faff, R. (2010). Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking & Finance, 34*(6), 1327-1343. doi:https://doi.org/10.1016/j.jbankfin.2009.11.028

Iyengar, S. (2010). Are sovereign credit ratings objective and transparent? *IUP Journal of Financial Economics, 8*(3), 7. Retrieved from https://www.proquest.com/openview/7ed9ce408bc2a38f07177c4f48b1c296/1?pq-origsite=gscholar&cbl=54458

Kabadayı, B., and Ahmet Alkan Çelik. (2015). Determinants of Sovereign Ratings in Emerging Countries: Qualitative Dependent Variables Panel Data Analysis. *International Journal of Economics and Financial Issues, 5*(3), 656-662. Retrieved from https://dergipark.org.tr/en/download/article-file/363044

Kabadayı, B., & Çelik, A. A. (2015). Determinants of Sovereign Ratings in Emerging Countries: Qualitative Dependent Variables Panel Data Analysis. *International Journal of Economics and Financial Issues, 5*(3), 656-662. Retrieved from https://dergipark.org.tr/en/download/article-file/363044

Kiff, J., Holland, A., Kisser, M., et al. (2010). The uses and abuses of sovereign credit ratings. *Global Financial Stability Report, 10*.

Kikulwe, E., & Asindu, M. (2020). Consumer demand and prospects for commercialization of nutritionally enhanced GM bananas in Uganda. *AgBioforum, 22*(1), 13-24. Retrieved from https://www.agbioforum.info/index.php/agb/article/view/18

Kräussl, R. (2005). Do credit rating agencies add to the dynamics of emerging market crises? *Journal of Financial Stability, 1*(3), 355-385. doi:https://doi.org/10.1016/j.jfs.2005.02.005

Kulkarni, A. D., & Lowe, B. (2016). Random forest algorithm for land cover classification. Retrieved from https://scholarworks.uttyler.edu/compsci_fac/1/

Kumar, K., & Haynes, J. D. (2003). Forecasting Credit Ratings Using An Ann And Statistical Techniques. *International journal of business studies, 11*(1). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.469.4256&rep=rep1&type=pdf

Kume, O. (2012). *Determinants of US corporate credit spreads*. Retrieved from https://rgu-repository.worktribe.com/output/248199

Mahomed Karodia, A., & Soni, D. (2014). South African economic woes: Poor political leadership and rating downgrades hampers growth and development. *Management Studies and Economic Systems, 1*(1), 51-66. Retrieved from https://academia-arabia.com/en/home/becomeamember?lang=en

Marandola, G., & Sinclair, T. J. (2017). Handbook on the Geographies of Money and Finance. In *Credit rating agencies are poorly understood and the rules developed for them will not work*: Edward Elgar Publishing. doi:https://doi.org/10.4337/9781784719005.00028.

Mathuria, M. (2013). Decision tree analysis on j48 algorithm for data mining. *Intrenational Journal ofAdvanced Research in Computer Science and Soft-ware Engineering, 3*(6).

Matthews, M., & Mokoena, B. A. (2020). The influence of service quality dimensions on customer satisfaction within visa facilitation centres in South Africa.

International Journal of eBusiness and eGovernment Studies, 12(2), 122-135. doi:https://doi.org/10.34111/ijebeg.20201220

Mellios, C., & Paget-Blanc, E. (2006). Which factors determine sovereign credit ratings? *The European Journal of Finance, 12*(4), 361-377. doi:https://doi.org/10.1080/13518470500377406

Meyer, D. F., & Mothibi, L. (2021). The Effect of Risk Rating Agencies Decisions on Economic Growth and Investment in a Developing Country: The Case of South Africa. *Journal of Risk and Financial Management, 14*(7), 288. doi:https://doi.org/10.3390/jrfm14070288

Mohamed, A. E. (2017). Comparative study of four supervised machine learning techniques for classification. *International Journal of Applied, 7*(2).

Mora, N. (2006). Sovereign credit ratings: Guilty beyond reasonable doubt? *Journal of Banking & Finance, 30*(7), 2041-2062. doi:https://doi.org/10.1016/j.jbankfin.2005.05.023

Moraes, V. Y. d., João Carlos Belloti, Fábio Ynoe de Moraes, José Antonio Galbiatti, Evandro Pereira Palácio, João Baptista Gomes dos Santos, and Flávio Faloppa. (2011). Hierarchy of evidence relating to hand surgery in Brazilian orthopedic journals. *Sao Paulo Medical Journal, 129*(2), 94-98. Retrieved from https://www.scielo.br/j/spmj/a/SWLX8gZBzZN4TLmrsGVgLWt/?format=pdf&lang=en

Mutize, M. a. N., M.P. (2021). International credit rating agencies in Africa: perceptions, trends and challenges. *International Journal of Sustainable Economy, 13*(1), 55-71. doi:https://doi.org/10.10.1504/ijse.2021.113303

Oskonbaeva, Z. (2020). Determinants of credit ratings: evidence from panel discrete model.

Ozturk, H., Namli, E., & Erdal, H. I. (2016). Modelling sovereign credit ratings: The accuracy of models in a heterogeneous sample. *Economic Modelling, 54*, 469-478. doi:https://doi.org/10.1016/j.econmod.2016.01.012

Pacelli, V. a. A., M. (2011). An artificial neural network approach for credit risk management". *Journal of Intelligent Learning Systems and Applications, 3*(2), 103. doi:https://doi.org/10.10.4236/jilsa.2011.32012

Polito, V. a. W., M.R. (2013). Sovereign credit ratings in the European Union: a model-based view". Mimeo.

Rai, B. (2017). Feature selection and predictive modeling of housing data using random forest. . *International Journal of Industrial and Systems Engineering, 11*(4), 5. Retrieved from https://publications.waset.org/10007029/feature-selection-and-predictive-modeling-of-housing-data-using-random-forest

Ratha, D., De, P. K., & Mohapatra, S. (2011). Shadow Sovereign Ratings for Unrated Developing Countries. *World Development, 39*(3), 295-307. doi:https://doi.org/10.1016/j.worlddev.2010.08.006

Rudd, M. P. H., GStat, J.M. and Priestley, J.L. (2017). A comparison of decision tree with logistic regression model for prediction of worst non-financial payment status in commercial credit. Retrieved from https://digitalcommons.kennesaw.edu/dataphdgreylit/5/

Schumacher, J., Trebesch, C., & Enderlein, H. (2021). Sovereign defaults in court. *Journal of International Economics, 131*, 103388. doi:https://doi.org/10.1016/j.jinteco.2020.103388

Sibanda, M. (2018). Determinants of sovereign credit ratings in an emerging economy: A case of South Africa (Doctoral dissertation, University of Cape Town).

Slabbert, A., Keeton, G. and Cattaneo, N. (2019). Investment-Grade or "Junk" Status: Do Sovereign Credit Ratings Really Matter.

Steinberg, D. a. C., N.S. . (1998). The hybrid CART-Logit model in classification and data mining. Salford Systems White Paper.The hybrid CART-Logit model in classification and data mining. Salford Systems White Paper. Retrieved from https://pdfs.semanticscholar.org/44da/cd9d4ff5fb495b30c8382b8267cb84f7d127.pdf

System, C. o. t. G. F. (2011a). Interactions of sovereign debt management with monetary conditions and financial stability", CGFS Papers, no 42, Basel.

Takawira, O., and Motseta, S.D. . (2021). The Relationship between Sovereign Credit Ratings and Capital Flows: A Case of South Africa. *International Journal of Economics and Finance Studies, 13*(1), pp. 251-276. doi:https://doi.org/10.10.34109/ijefs.202112234

Takawira, O., & Mwamba, J. W. M. (2021). The effects of sovereign ratings on financial stability in South Africa. *African Journal of Business and Economic Research, 16*(1), 27-65. doi:https://doi.org/10.10.31920/1750-4562/2021/v16n1a2

Takawira, O. a. M., W.M. (2020). Determinants of Sovereign Credit Ratings: An Application of the Naïve Bayes Classifier. *Eurasian Journal of Economics and Finance., 8*(4), 279-299. doi:https://doi.org/10.10.15604/ejef.2020.08.04.008

Tangirala, S. (2020). Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm. . *International Journal of Advanced Computer Science and Applications, , 11*(2), 612-619. Retrieved from https://pdfs.semanticscholar.org/3679/7533a57bc76b376129aa258fcc426d94d0c1.pdf

Valencia, G. A. D. (2020). Forms of informal financing of informal traders in Colombia Cases: Cúcuta, Ibagué and Villavicencio. *Cuadernos de Economía, 43*(123), 259-274. Retrieved from http://cude.info/index.php/CUDE/article/view/2

Wallenius, C., Alvinius, A., & Larsson, G. (2020). Decision-Making in a Military Staff Context. *Res Militaris, 10*(1), 1-17. Retrieved from https://www.researchgate.net/profile/Claes-Wallenius/publication/358590436

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Weyers, K., & Eliott, A. (2017). Checkmate : South Africa's credit rating downgraded to “junk” status. *Without Prejudice, 17*(4), 10-13. doi:https://doi.org/10.10.10520/EJC-7fe320f11

Yoon, K. P., Sedaghat, M., & Kim, G. (2020). Portfolio Selection by the Axiom of Choice: Post Mean-Variance Analysis. *International Journal of Operations and Quantitative Management, 26*(4), 303-318. Retrieved from https://web.archive.org/web/20210627140241id_/http://www.ijoqm.org/papers/26-4-4-p.pdf

Zacharis, N. Z. (2018). Classification and regression trees (CART) for predictive modelling in blended learning. IJ Intelligent Systems and Applications,. *3*, 1-9. doi:https://doi.org/10.10.5815/ijisa.2018.03.0

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 28-87) **Doi:** 10.34109/ijefs.202112222

APPENDIX

Appendix A: Descriptive Statistics

The descriptive statistics results are presented in the table below:

Table A1: Descriptive Statistics

Statistic	FDGDP	REER	PIR	BOP	CAB	UR	CPIH	GDPpc	HDDI
Mean	12.5945	87.6733	11.7802	8367.950	-22080.81	25.0200	68.5650	2.7550	70.8863
Median	12.1680	88.1567	10.5000	12351.00	-17130.50	25.0000	65.5500	3.0000	75.3500
Maximum	22.6443	104.1000	21.0000	93797.00	6771.000	29.3000	109.4667	7.1000	87.8000
Minimum	6.0850	67.3233	8.5000	-109916.0	-72744.00	21.0000	40.2209	-2.6000	51.7000
Std. deviation	3.2211	8.3873	2.8000	44430.83	20335.78	1.8877	20.6579	1.8764	11.5431
Skewness	1.0024	-.1632	.9306	-.4385	-.5983	.1415	.4225	-.2749	-.4502
Kurtosis	4.3452	2.4572	3.2208	2.6162	2.3898	2.4463	1.9309	3.1076	1.6639
Jarque-Bera	19.4287	1.3372	11.7092	3.0549	6.0142	1.2891	6.1899	1.0465	8.6525
Probability	.001	.5124	.0029	.2171	.0494	.5249	.0453	.5926	.0132

Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 28-87) **Doi:** 10.34109/ijefs.202112222

Appendix C: Trees Size and Mtry

Figure C.1

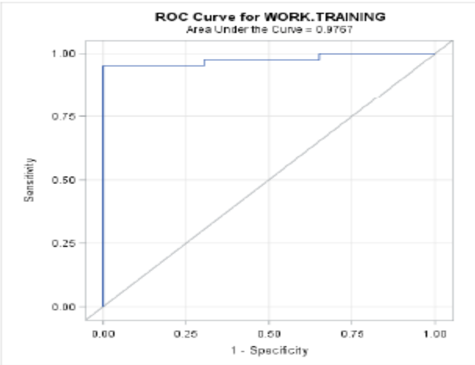


Figure a: ROC Curve for Fitch Train Data

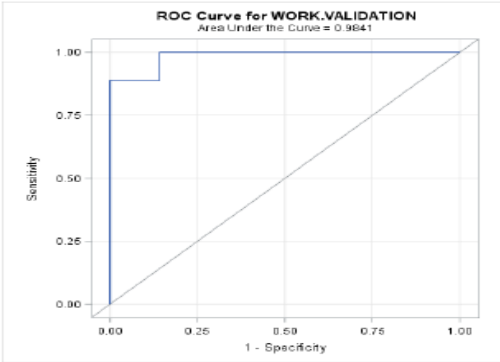


Figure C: ROC Curve for Fitch Test Data

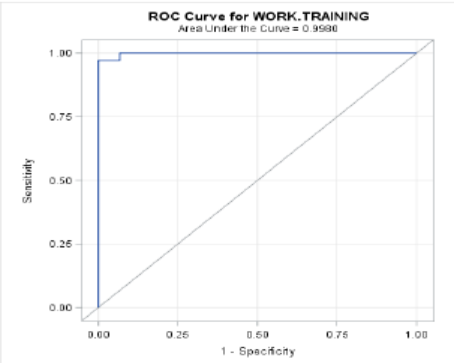


Figure D: ROC Curve for Moody's Train Data

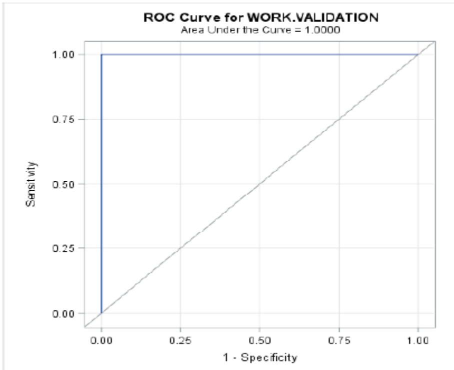


Figure E: ROC Curve for Moody's Test Data

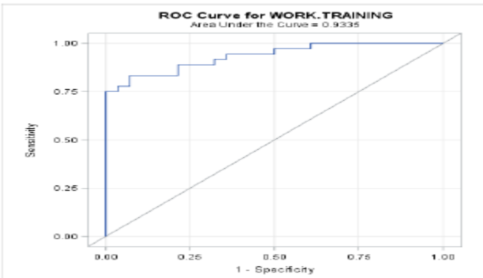


Figure e: ROC Curve for SNPoor Train Data

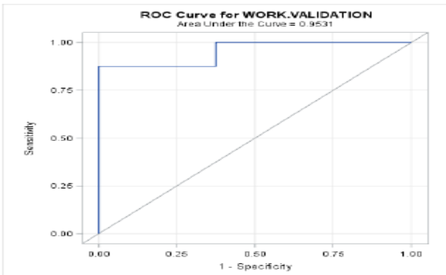


Figure f: ROC Curve for SNPoor Test Data

Figure C.1 ROC Curves (Source: By Author)

INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Figure C.2

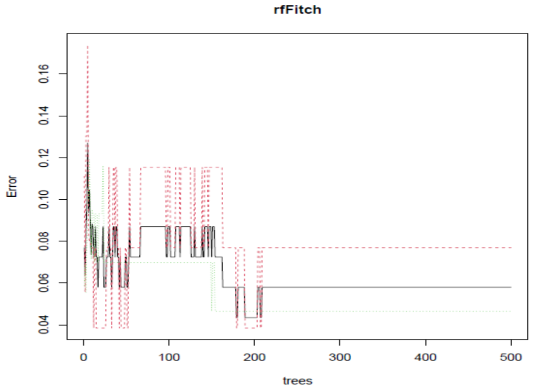


Figure (a) Effect of tree size to error for Fitch

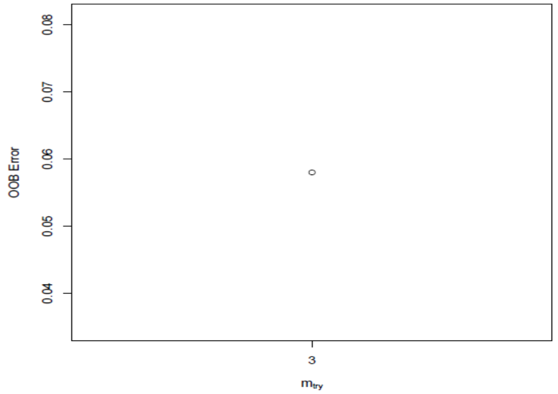


Figure (b) Mtry best parameter for Fitch

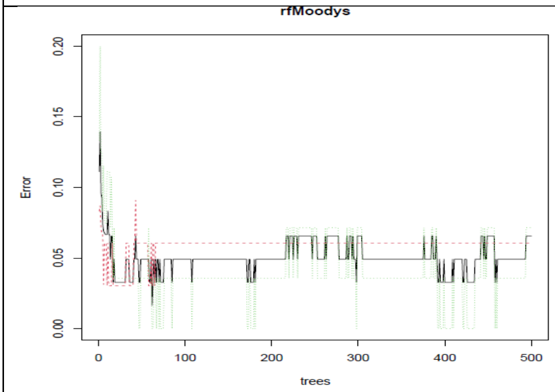


Figure (c) Effect of tree size to error for Moody

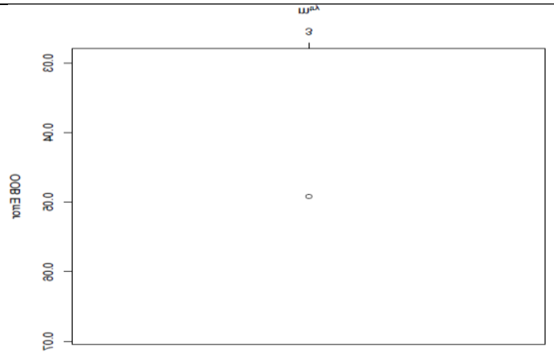


Figure (b) Mtry best parameter for Fitch

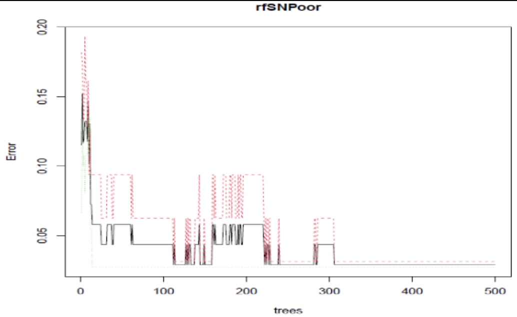


Figure (e) Effect of treesize to error for SNPoor

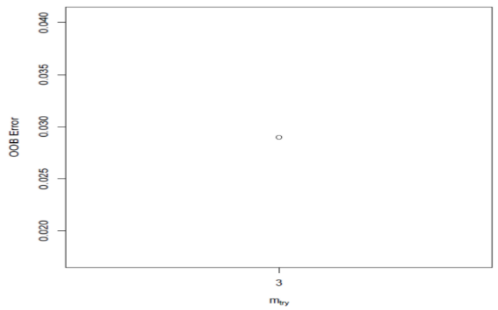


Figure (f) Mtry best parameter for SNPoor

Figure C.2 Trees Size and Mtry
Source: By Author

Figure C.2 indicated the best parameters to optimise the model.

Appendix D: Partial Dependence Plots

The partial plots for the three important variables are shown in Figure D1 to Figure D3.

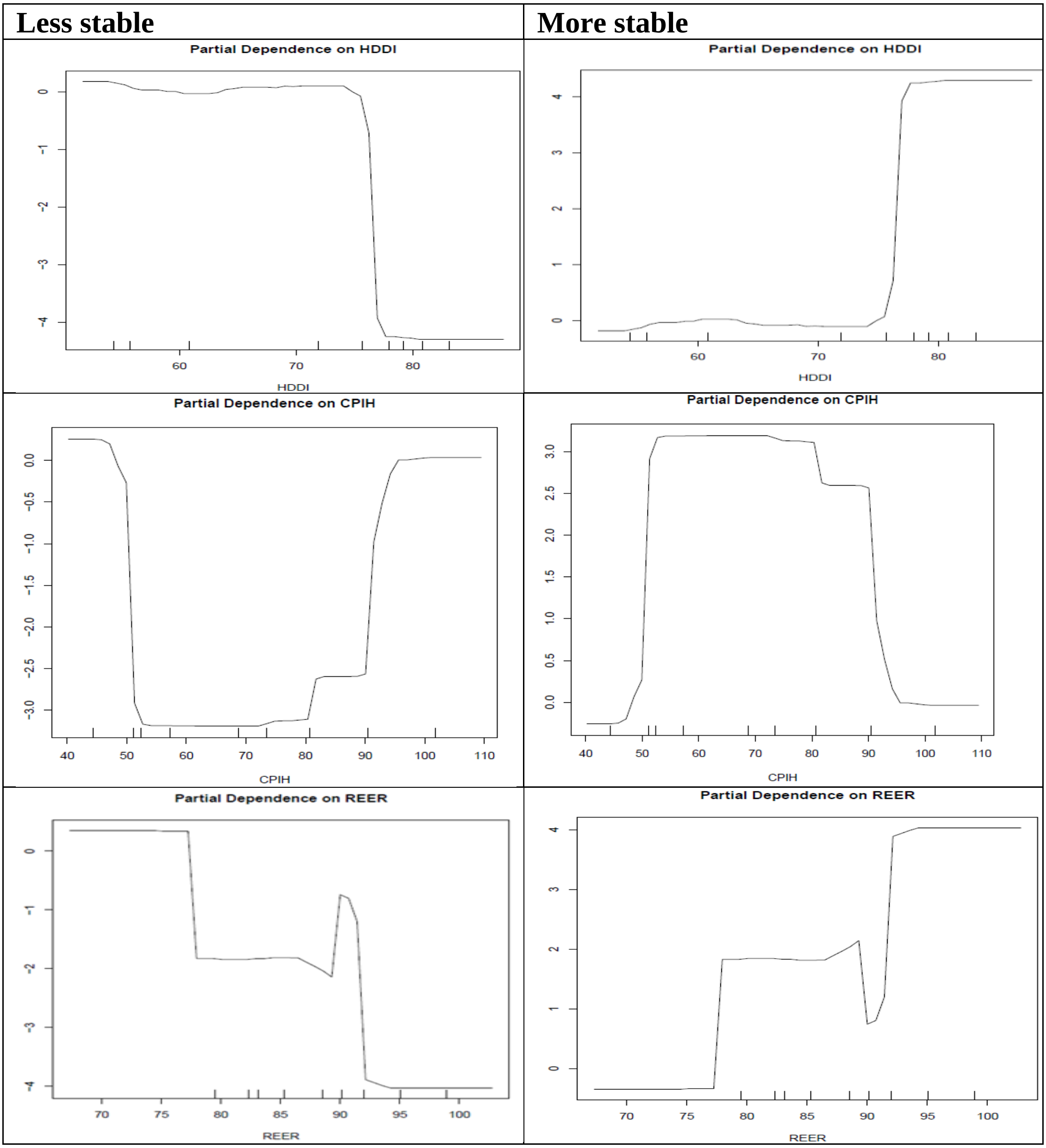


Figure D1: Partial dependence model for Fitch
Source: By Author

The partial plots for the three important variables for Moody's are shown in Figure D2.

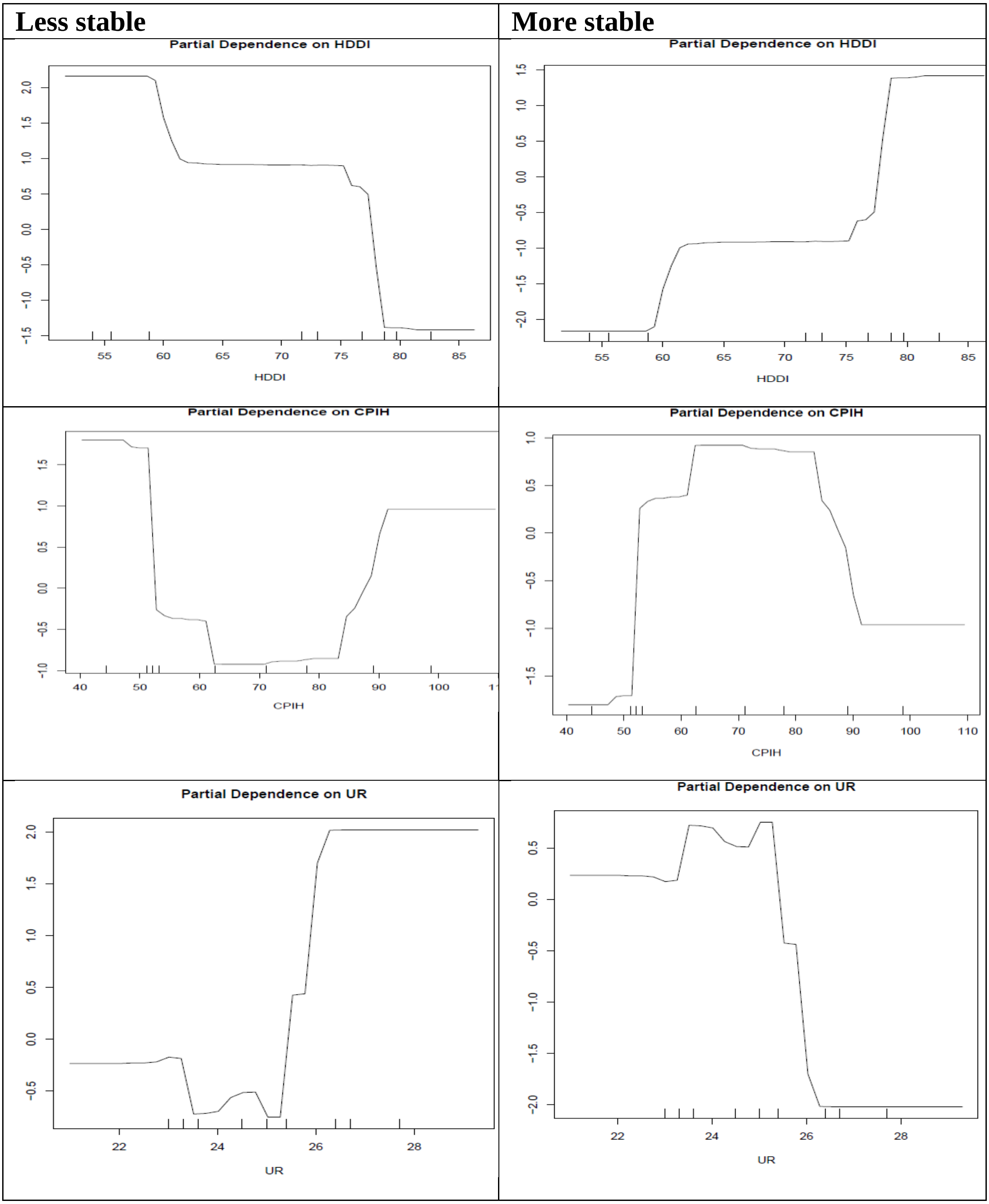


Figure D2: Partial dependence model for Moody's
Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

The partial plots for the three important variables for SNPoor are shown in Figure D3.

Less stable	**More stable**

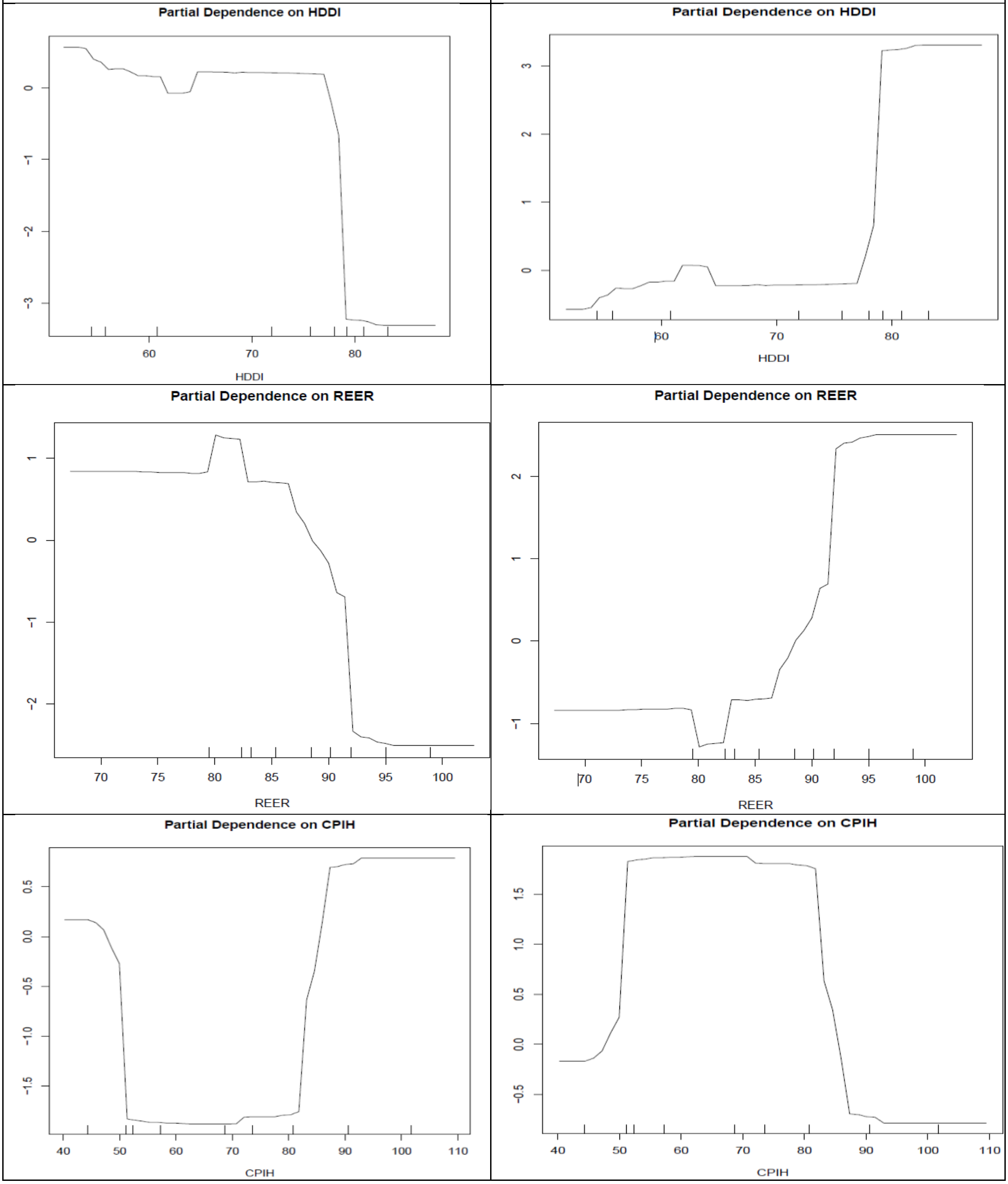


Figure D3: Partial dependence model for SNPoor
Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Appendix E: Random Forest Prediction Trees

<u>Prediction using Fitch model Continued</u>

Figure E1 below shows the predicted class for Fitch using tree two.

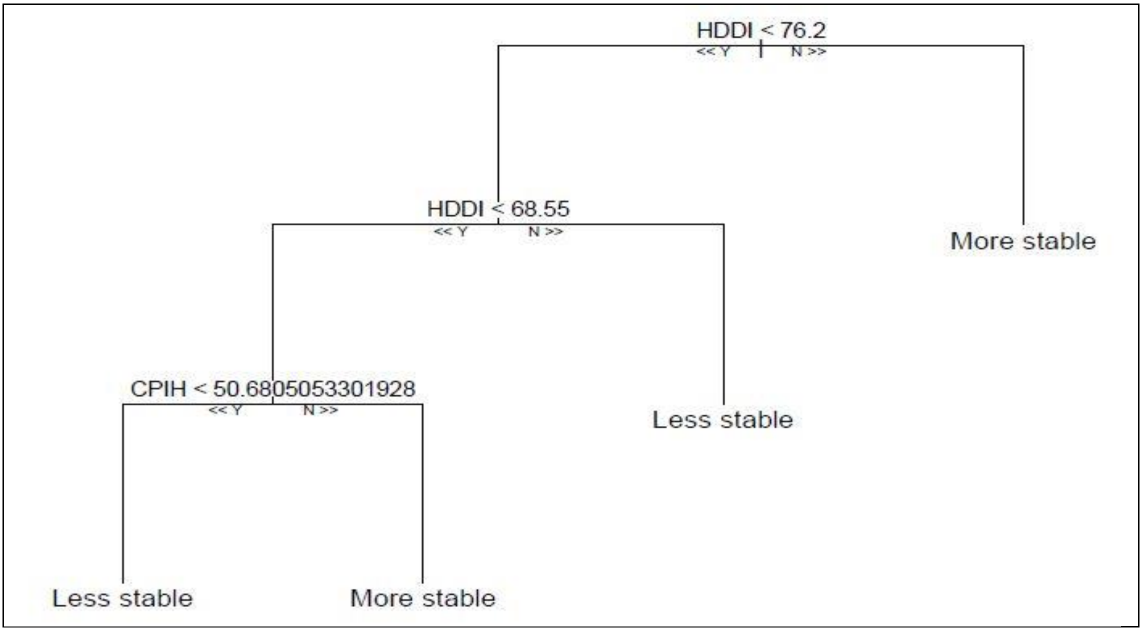


Figure E1: Prediction tree for Fitch Model Tree 2
Source: By Author

Fitch model tree 2 had four paths using the variables HDDI and CPIH.

Path 1: If HDDI is less than 76.2, HDDI is less than 68.55 and CPIH is less than 50.6805, then the classification is less stable.

Path 2: If HDDI is less than 76.2, HDDI is less than 68.55 and CPIH is more than 50.6805, then the classification is more stable.

Path 3: If HDDI is less than 76.2 and HDDI is more than 68.55, the classification is less stable.

Path 4: If HDDI is more than 76.2, the classification is more stable.

Observations 1 to 5 had HDDI less than 76.2 and more than 68.55 and were classified as less stable. The last observation had HDDI of more than 76.2 and was classified as more stable, and the predicted class and the current class are shown in Table EE1.

Vol: 14 No: 1 Year: 2022 ISSN: 1309-8055 (Online) (pp. 29-87) Doi: 10.34109/ijefs. 202220002

Table EE1: Fitch Prediction Model for Tree 2

Time	HDDI	CPIH	Fitch Class	Predicted Fitch Class Tree 1
2019/03/31	73	110.1	0	0
2019/06/30	72.9s	112.0333	0	0
2019/09/30	72.6	113.1	0	0
2019/12/31	73.2	113.5667	0	0
2020/03/31	73.6	114.9667	0	0
2020/06/30	85.3	114.7333	0	1

Source: On Author

The second tree model for Fitch correctly predicted the classification of five out of the six observations. Only the model classified the last observation as more stable (“1”).

Figure E2 below shows the predicted class for Fitch using tree three.

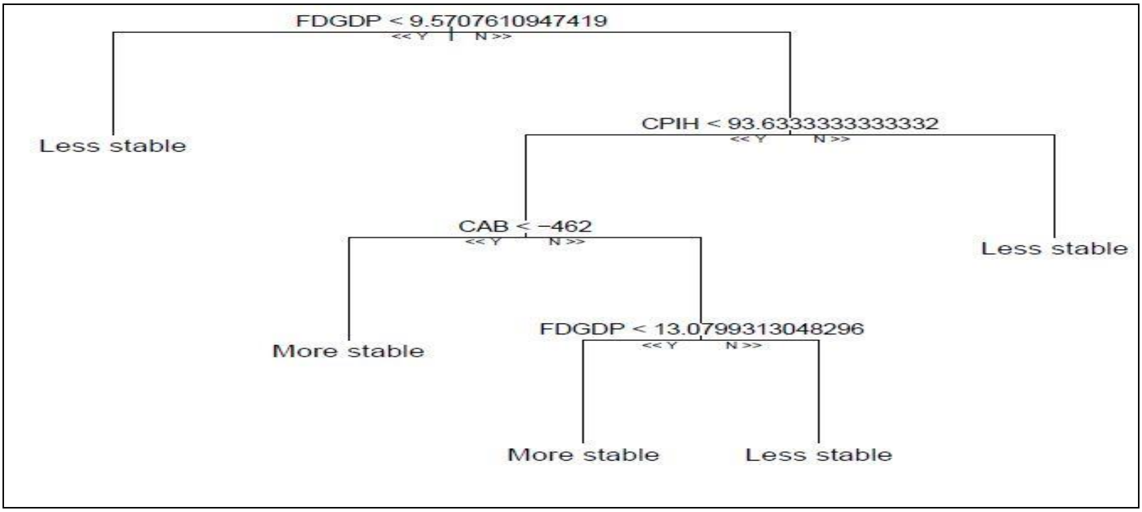


Figure E2: Prediction tree for Fitch Model Tree 3
Source: By Author

The classification of the future observations for Fitch model tree 3 was done using the variables FDGDP, CPIH, and CAB using five paths.

Path 1: If FDGDP is less than 9.5708, the observation is less stable.

Path 2: If FDGDP is more than 9.5708, CPIH is less than 93.6333, and CAB is less than -462, then the classification is more stable.

Path 3: If FDGDP is more than 9.5708, CPIH is less than 93.6333, CAB is more than -462, and FDGDP is less than 13.0799, then the classification is more stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Path 4: If FDGDP is more than 9.5708, CPIH is less than 93.6333, CAB is more than -462, and FDGDP is more than 13.0799, then the classification is less stable.

Path 5: If FDGDP is more than 9.5708 and CPIH is more than 93.6333, then the classification is less stable.

All the observations had FDGDP more than 9.5708 and CPIH more than 93.6333 and were classified as less stable. The predicted class and the current class are shown in Table EE2.

Table EE2: Fitch Prediction Model for Tree 3

Time	FDGDP	CPIH	CAB	Fitch Class	Predicted Fitch Class Tree 1
2019/03/31	17.045853	110.1	-57063.9368	0	0
2019/06/30	16.048351	112.0333	-29185.8365	0	0
2019/09/30	17.485446	113.1	-62178.4028	0	0
2019/12/31	19.663271	113.5667	57.3302	0	0
2020/03/31	21.027019	114.9667	-8168.3622	0	0
2020/06/30	26.170115	114.7333	-8503.0976	0	0

Source: On Author

The third tree model for Fitch correctly predicted the classification of all the observations, that is, the predicted class was less stable (“0”).

Figure E3 below shows the predicted class for Fitch using tree four.

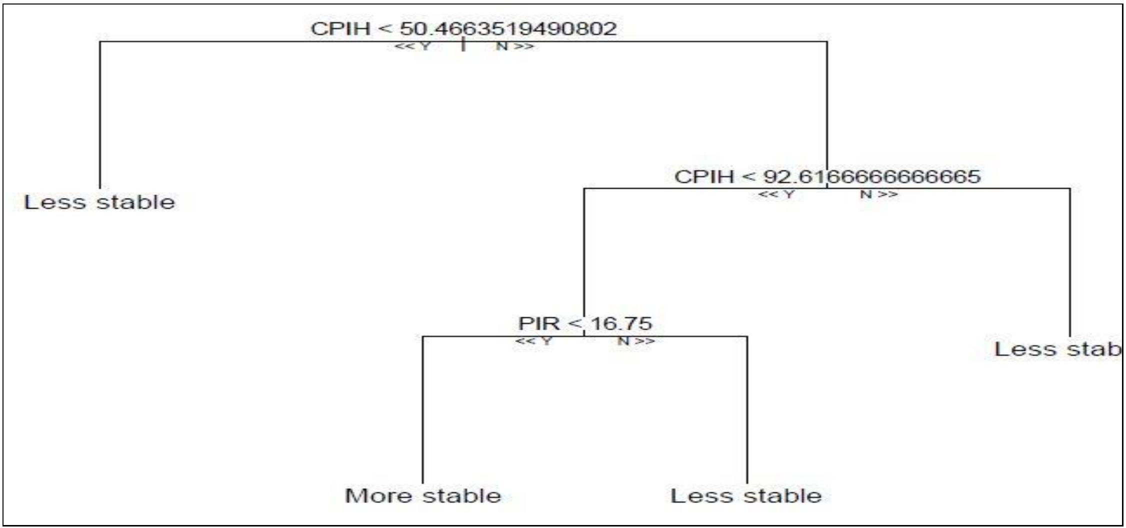


Figure E3: Prediction tree for Fitch Model Tree 4
Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

The classification of the future observations for Fitch model tree 4 was done using the variables CPIH and PIR using four paths.

Path 1: If CPIH is less than 50.4664, then the observation is classified as less stable.

Path 2: If CPIH is greater than 50.4664, CPIH is less than 92.6167 and PIR is less than 16.75, then the classification is more stable.

Path 3: If CPIH is greater than 50.4664, CPIH is less than 92.6167, and PIR is more than 16.75, then the classification is less stable.

Path 4: If CPIH is greater than 50.4664 and CPIH is more than 92.6167, then the classification is less stable.

All the observations had CPIH greater than 50.4664 and CPIH more than 92.6167 and were classified as less stable. Using the tree generated in Figure E3, the predicted class and the current class are shown in Table EE3.

Table EE3: Fitch Prediction Model for Tree 4

Time	**CPIH**	**PIR**	**Fitch Class**	**Predicted Fitch Class Tree 1**
2019/03/31	110.1	10.25	0	0
2019/06/30	112.03333	10.25	0	0
2019/09/30	113.1	10	0	0
2019/12/31	113.56667	10	0	0
2020/03/31	114.96667	9.416667	0	0
2020/06/30	114.73333	7.416667	0	0

Source: On Author

The fourth tree model for Fitch correctly predicted the classification of all the observations. That is, the predicted class was less stable (“0”).

Figure E4 below shows the predicted class for Fitch using tree five.

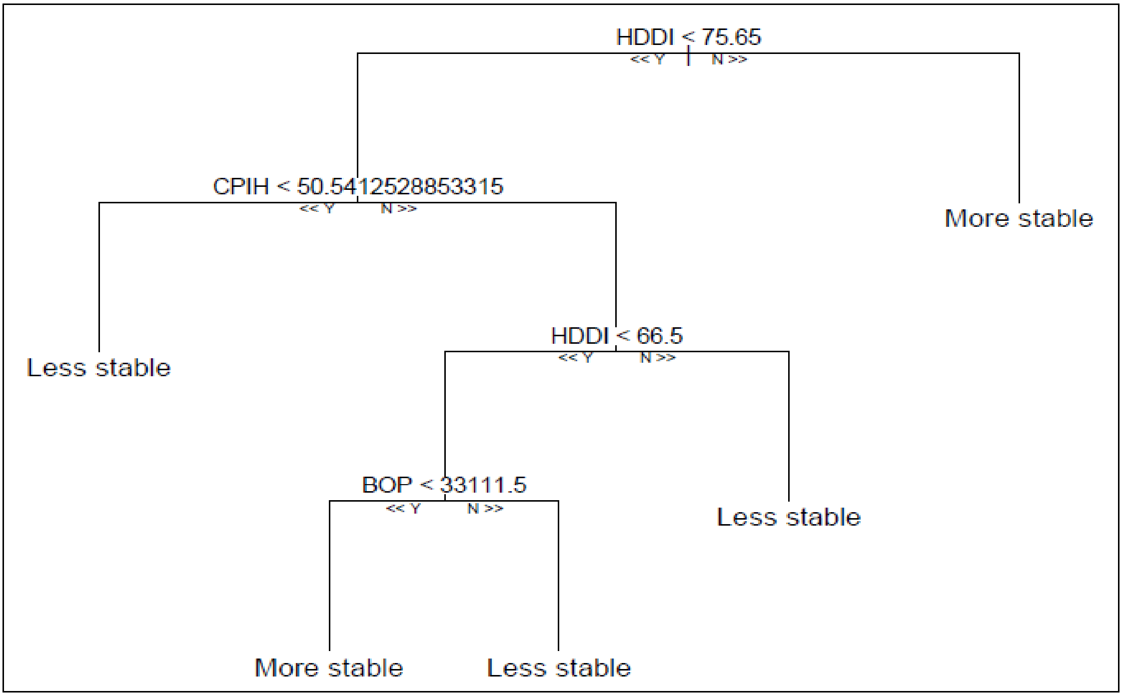


Figure E4: Prediction tree for Fitch Model Tree 5
Source: By Author

Fitch model tree 5 had five paths using HDDI, CPIH, and BOP.

Path 1: If HDDI is less than 75.65 and CPIH is less than 50.5413, the classification is less stable.

Path 2: If HDDI is less than 75.65, CPIH is more than 50.5413, HDDI is less than 66.5, and BOP is less than 33.1115, the classification is more stable.

Path 3: If HDDI is less than 75.65, CPIH is more than 50.5413, HDDI is less than 66.5, and BOP is more than 33.1115, the classification is less stable.

Path 4: If HDDI is less than 75.65, CPIH is more than 50.5413, and HDDI is more than 66.5, then the classification is less stable

Path 4: If HDDI is more than 75.65, the classification is more stable.

Observations 1 to 5 had HDDI less than 75.65, CPIH more than 50.5413, and HDDI more than 66.5 and were classified as less stable. The last observation had HDDI of more than 75.65 and was classified as more stable. Using the tree generated in Figure E4, the predicted class and the current class are shown in Table EE4.

Vol: 14 No: 1 Year: 2022 ISSN: 1309-8055 (Online) (pp. 29-87) Doi: 10.34109/ijefs. 202220002

Table EE4: Fitch Prediction Model for Tree 5

Time	HDDI	CPIH	BOP	Fitch Class	Predicted Fitch Class Tree 1
2019/03/31	73	110.1	42301	0	0
2019/06/30	72.9	112.0333	-31495	0	0
2019/09/30	72.6	113.1	43992	0	0
2019/12/31	73.2	113.5667	102486	0	0
2020/03/31	73.6	114.9667	201663	0	0
2020/06/30	85.3	114.7333	91508	0	1

Source: Author

The fifth tree model for Fitch correctly predicted the classification of five out of the six observations, only the last observation was classified as more stable by the model (“1”).

Figure E5 below shows the predicted class for Fitch using tree six.

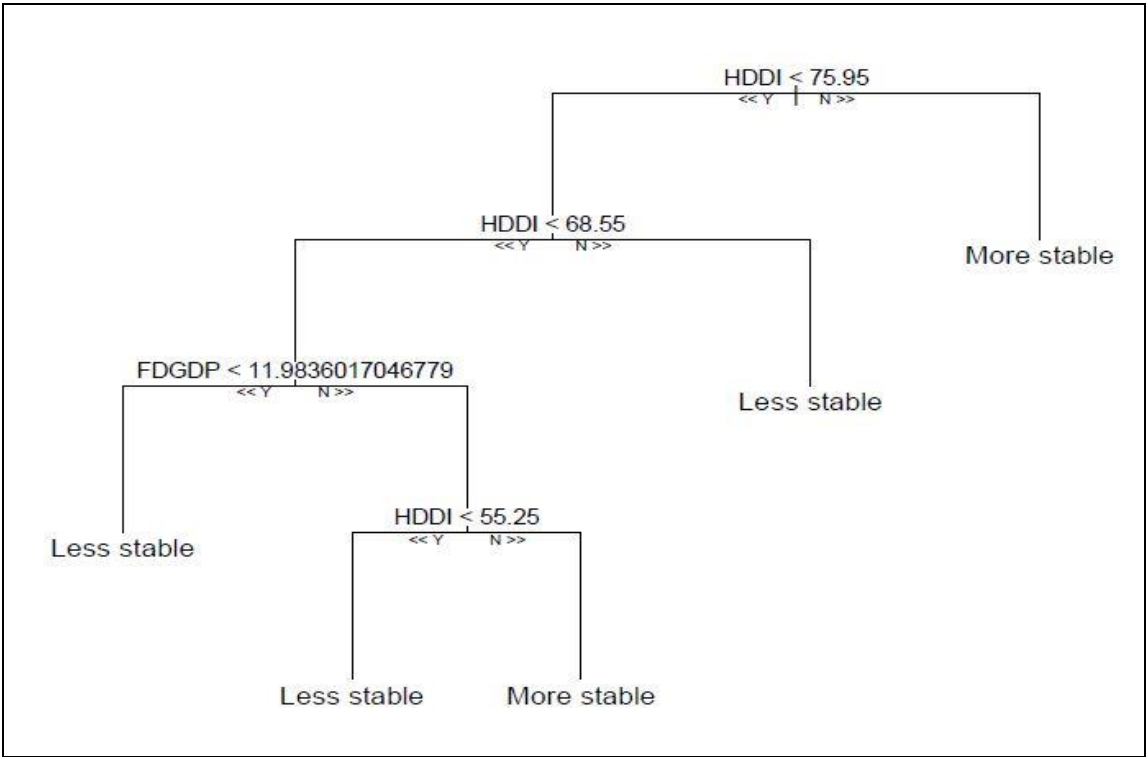


Figure E5: Prediction tree for Fitch Model Tree 6
Source: By Author

Fitch model tree 6 had five paths using the variables HDDI and FDGDP.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Path 1: If HDDI is less than 75.95, HDDI is less than 68.55, and FDGDP is less than 11.9836, the classification is less stable.

Path 2: If HDDI is less than 75.95, HDDI is less than 68.55, FDGDP is more than 11.9836, and HDDI is less than 55.25, then the classification is less stable.

Path 3: If HDDI is less than 75.95, HDDI is less than 68.55, FDGDP is more than 11.9836, and HDDI is more than 55.25, the classification is more stable.

Path 4: If HDDI is less than 75.95 and HDDI is more than 68.55, the classification is less stable.

Path 4: If HDDI is more than 75.95, the classification is more stable.

Observations 1 to 5 had HDDI less than 75.65 and HDDI more than 68.55; then the classifications were less stable. The last observation had HDDI of more than 75.95 and was classified as more stable. Using the tree generated in Figure E5, the predicted class and the current class are shown in Table EE5.

Table EE5: Fitch Prediction Model for Tree 6

Time	HDDI	FDGDP	Fitch Class	Predicted Fitch Class Tree 1
2019/03/31	73	17.04585	0	0
2019/06/30	72.9	16.04835	0	0
2019/09/30	72.6	17.48545	0	0
2019/12/31	73.2	19.66327	0	0
2020/03/31	73.6	21.02702	0	0
2020/06/30	85.3	26.17012	0	1

Source: On Author

The sixth tree model for Fitch correctly predicted the classification of five out of the six observations. Only the model classified the last observation as more stable (“1”).

<u>Prediction using Moodys model Continued</u>

Figure E6 below shows the predicted class for Moody’s using tree two.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

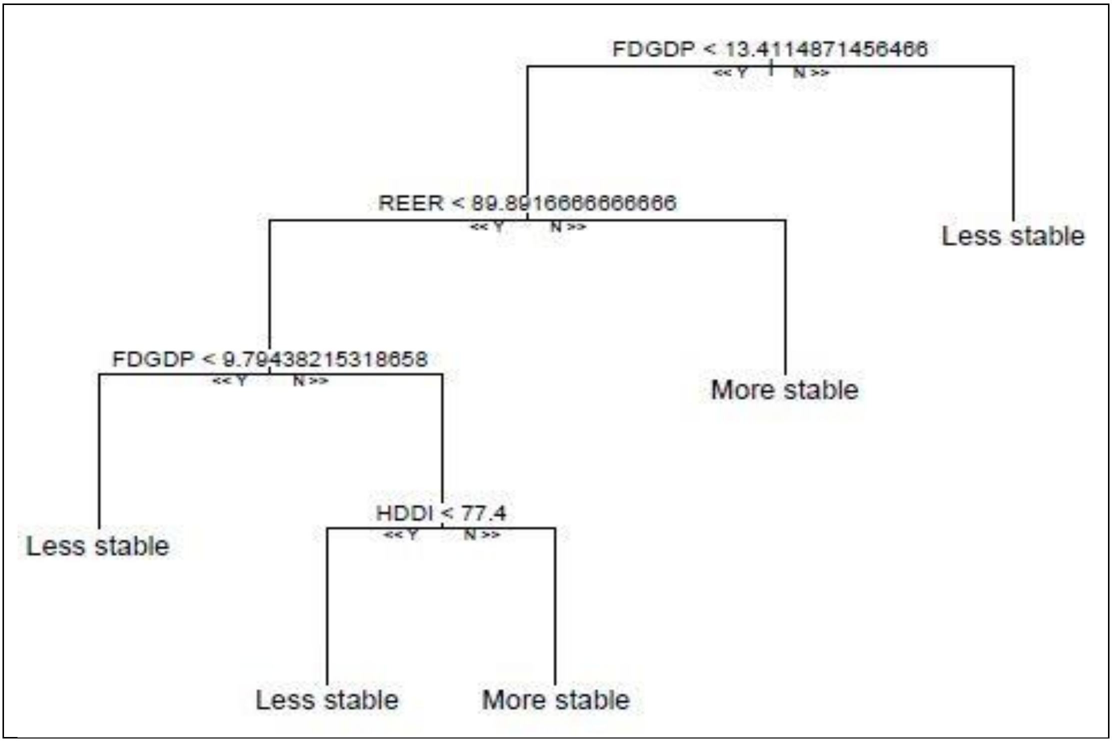


Figure E6: Prediction tree for Moodys Model Tree 2

Source: By Author

The future observations for Moody's model tree 2 were classified using the variables FDGDP, REER, and HDDI using five paths.

Path 1: If FDGDP is less than 13.4115, REER is less than 89.8917, and FDGDP is less than 9.7944, the observation is classified as less stable.

Path 2: If FDGDP is less than 13.4115, REER is less than 89.8917, FDGDP is more than 9.7944, and HDDI is less than 77.4, the observation is classified as less stable.

Path 3: If FDGDP is less than 13.4115, REER is less than 89.8917, FDGDP is more than 9.7944, and HDDI is more than 77.4, the observation is classified as more stable.

Path 4: If FDGDP is less than 13.4115 and REER is more than 89.8917, the observation is classified as more stable.

Path 5: If FDGDP is more than 13.4115, the classification is less stable.

All the observations had FDGDP, more than 13.4115, and were classified as less stable. Using the tree generated in Figure E6, the predicted class and the current class are shown in Table EE6.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Table EE6: Moody's Prediction Model for Tree 2

Time	FDGDP	REER	HDDI	Moodys Class	Predicted Moodys Class Tree 2
2019/03/31	17.04585	93.57	73	0	0
2019/06/30	16.04835	92.01333	72.9	0	0
2019/09/30	17.48545	91.33667	72.6	0	0
2019/12/31	19.66327	91.28667	73.2	0	0
2020/03/31	21.02702	88.19333	73.6	0	0
2020/06/30	26.17012	76.50333	85.3	0	0

Source: On Author

The second tree model for Moody’s correctly predicted all the observations’ classification; that is, the predicted class was less stable (“0”).

Figure E7 below shows the predicted class for Moody’s using tree three.

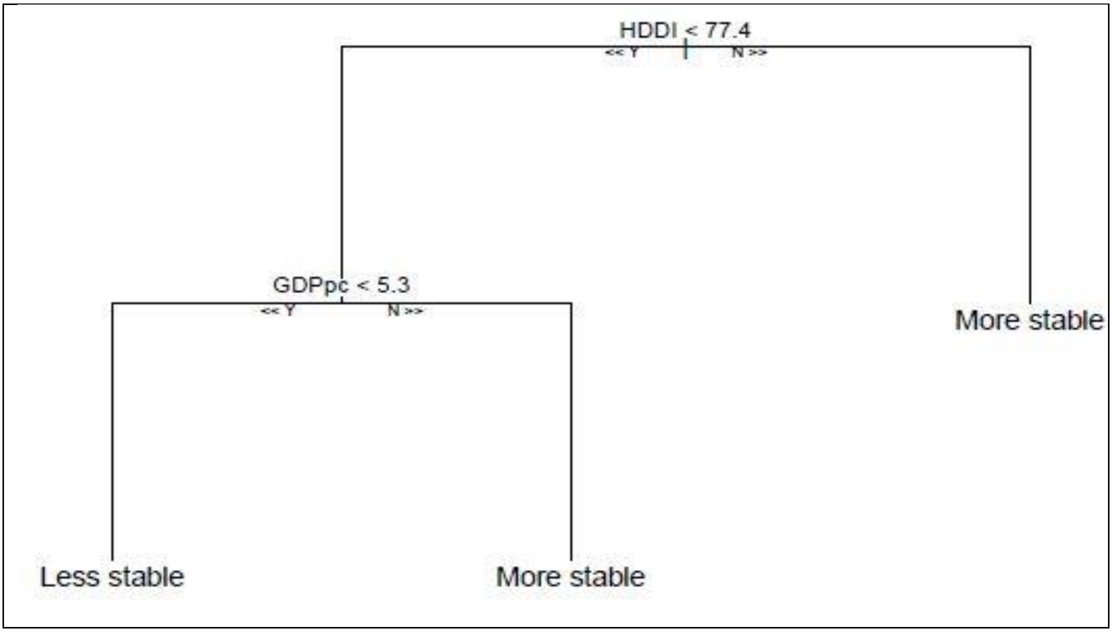


Figure E7: Prediction tree for Moodys Model Tree 3
Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

The future observations for Moody's model tree 3 were classified using the variables HDDI and GDPpc through three paths.

Path 1: If HDDI is less than 77.4 and GDPpc is less than 5.3, the observation is classified as less stable.

Path 2: If HDDI is less than 77.4 and GDPpc is more than 5.3, the observation is classified as more stable.

Path 3: If HDDI is more than 77.4, the observation is classified as more stable.

Observations 1 to 5 had HDDI less than 77.4 and GDPpc less than 5.3 and were classified as less stable; the last observation had HDDI more than 77.4 and was classified as more stable. Using the tree generated in Figure E7, the predicted class and the current class are shown in Table EE7.

Table EE7: Moody's Prediction Model for Tree 3

Time	HDDI	GDPpc	Moodys Class	Predicted Moodys Class Tree 3
2019/03/31	73	0	0	0
2019/06/30	72.9	0.9	0	0
2019/09/30	72.6	0.1	0	0
2019/12/31	73.2	-0.5	0	0
2020/03/31	73.6	0.1	0	0
2020/06/30	85.3	-17.1	0	1

Source: On Author

The third tree model for Moody's correctly predicted the classification of five out of the six observations, only the last observation was classified as more stable by the model ("1").

Figure E8 below shows the predicted class for Moodys using tree four.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

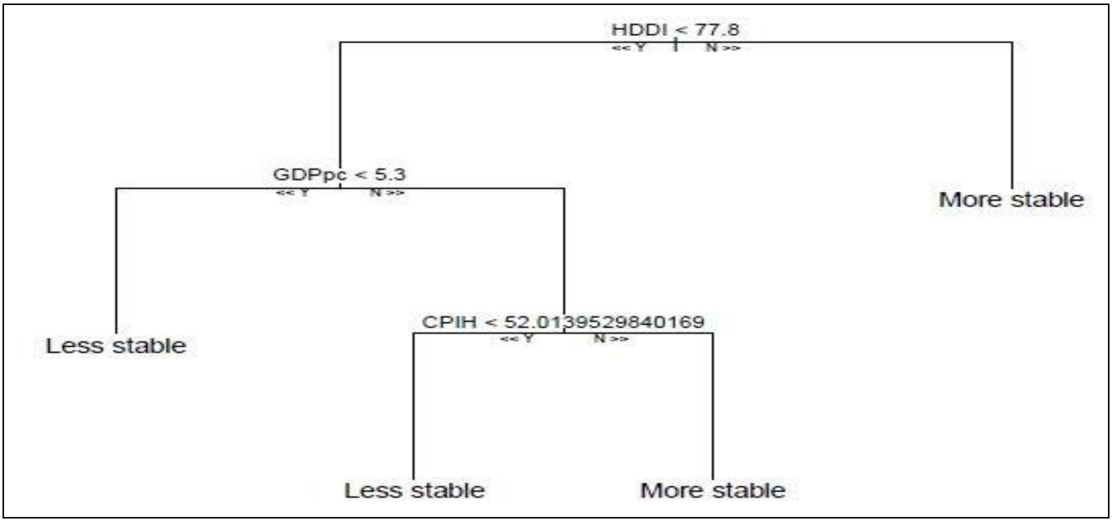


Figure E8: Prediction tree for Moody's Model Tree 4
Source: By Author

Moody's model tree 4 was constructed using four paths: HDDI, GDPpc, and CPIH.

Path 1: If HDDI is less than 77.8 and GDPpc is less than 5.3, the observation is classified as less stable.

Path 2: If HDDI is less than 77.8, GDPpc is more than 5.3 and CPIH is less than 52.0140, the observation is less stable.

Path 3: If HDDI is less than 77.8, GDPpc is more than 5.3 and CPIH is more than 52.0140, the observation is more stable.

Path 4: If HDDI is more than 77.8, the observation is classified as more stable.

Observations 1 to 5 had HDDI less than 77.8 and GDPpc less than 5.3 and were classified as less stable, and the last observation had HDDI more than 77.8 and was classified as more stable. Using the tree generated in Figure E8, the predicted class and the current class are shown in Table EE8.

Table EE8: Moody's Prediction Model for Tree 4

Time	HDDI	GDPpc	CPIH	Moodys Class	Predicted Moodys Class Tree 4
2019/03/31	73	0	110.1	0	0
2019/06/30	72.9	0.9	112.0333333	0	0
2019/09/30	72.6	0.1	113.1	0	0
2019/12/31	73.2	-0.5	113.5666667	0	0

2020/03/31	73.6	0.1	114.9666667	0	0
2020/06/30	85.3	-17.1	114.7333333	0	1

Source: On Author

The fourth tree model for Moody's correctly predicted all the observations' classification; that is, the predicted class was less stable ("0").

Figure E9 below shows the predicted class for Moody's using tree five.

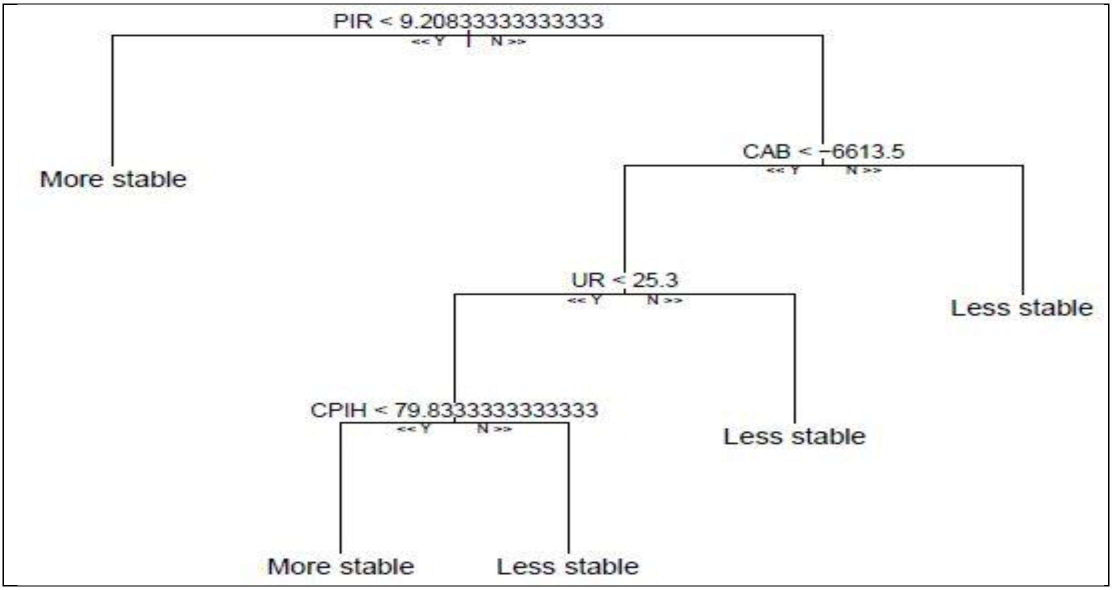


Figure E9: Prediction tree for Moodys Model Tree 5
Source: By Author

Moody's model tree 5 was constructed using five paths: PIR, CAB, UR, and CPIH.

Path 1: If PIR is less than 9.2083, the observation is classified as more stable.

Path 2: If PIR is more than 9.2083, CAB is less than -6613.5, UR is less than 25.3, and CPIH is less than 79.8333, the observation is classified as more stable.

Path 3: If PIR is more than 9.2083, CAB is less than -6613.5, UR is less than 25.3, and CPIH is more than 79.8333, the observation is classified as less stable.

Path 4: If PIR is more than 9.2083, CAB is less than -6613.5, UR is more than 25.3, then the observation is classified as less stable.

Path 5: If PIR is more than 9.2083 and CAB is more than -6613.5, the observation is classified as less stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Observations 1, 2, 3, and 5 had PIR more than 9.2083, CAB less than -6613.5, and UR more than 25.3, and then the observations were classified as less stable. Observations 4 had PIR of more than 9.2083, CAB more than -6613.5 and was classified as less stable. The last observation had PIR of less than 9.2083 and was classified as more stable. Using the tree generated in Figure E9, the predicted class and the current class are shown in Table EE9.

Table EE9: Moody's Prediction Model for Tree 5

Time	PIR	CAB	UR	CPIH	Moodys Class	Predicted Moodys Class Tree 5
2019/03/31	10.25	-57063.9	27.6	110.1	0	0
2019/06/30	10.25	-29185.8	29	112.0333333	0	0
2019/09/30	10	-62178.4	29.1	113.1	0	0
2019/12/31	10	57.3302	29.1	113.5666667	0	0
2020/03/31	9.416667	-8168.36	30.1	114.9666667	0	0
2020/06/30	7.416667	-8503.1	23.3	114.7333333	0	1

Source: Author

The fifth tree model for Moody's correctly predicted the classification of five out of the six observations. Only the model classified the last observation as more stable ("1").

Figure E10 below shows the predicted class for Moody's using tree six.

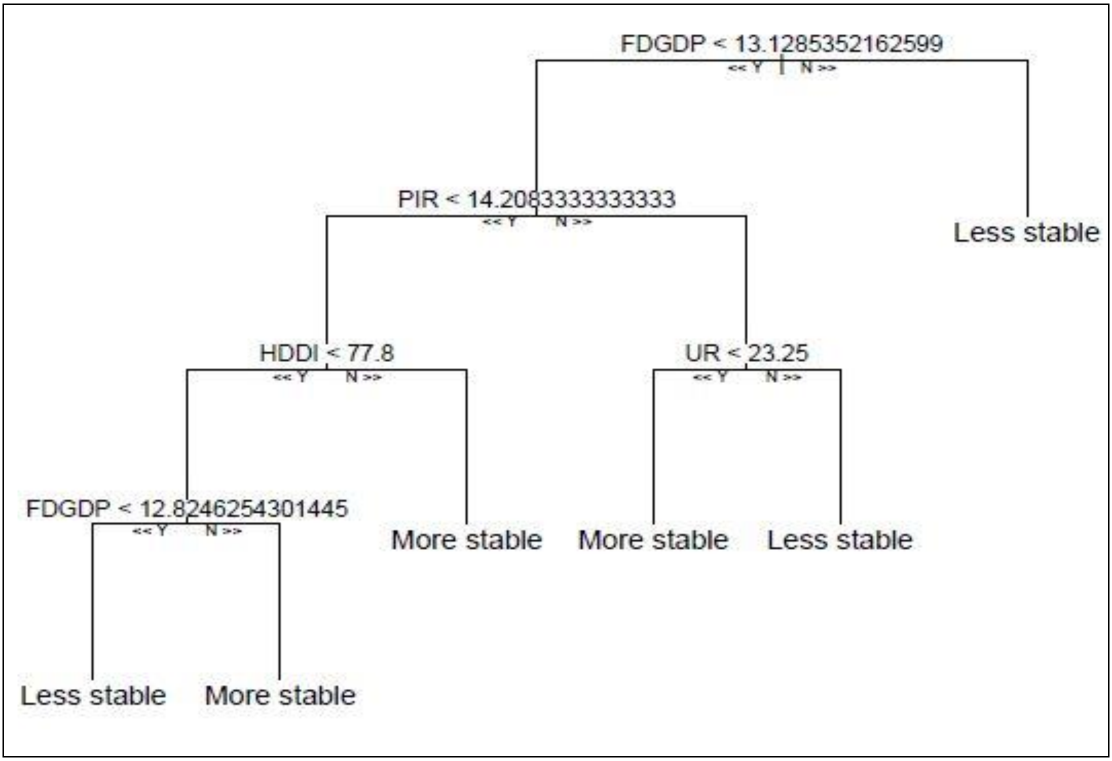


Figure E10: Prediction tree for Moody's Model Tree 6
Source: By Author

The future observations for Moody's model tree 6 were classified using the variables FDGDP, PIR, HDDI, and UR using six paths.

Path 1: If FDGDP is less than 13.1285, PIR is less than 14.2083, HDDI is less than 77.8, and FDGDP is less than 12.8246, the observation is classified as less stable.

Path 2: If FDGDP is less than 13.1285, PIR is less than 14.2083, HDDI is less than 77.8, and FDGDP is more than 12.8246, the observation is classified as more stable.

Path 3: If FDGDP is less than 13.1285, PIR is less than 14.2083, and HDDI is more than 77.8, then the observation is classified as more stable.

Path 4: If FDGDP is less than 13.1285, PIR is more than 14.2083, and UR is less than 23.25, the classification is more stable.

Path 5: If FDGDP is less than 13.1285, PIR is more than 14.2083, and UR is more than 23.25, the classification is less stable.

Path 6: If FDGDP is more than 13.1285, the classification is less stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

All the observations had FDGDP of more than 13.1285 and were classified as less stable. Using the tree generated in Figure E10, the predicted class and the current class are shown in Table EE10.

Table EE10: Moody's Prediction Model for Tree 6

Time	FDGDP	PIR	HDDI	UR	Moodys Class	Predicted Moodys Class Tree 5
2019/03/31	17.04585	10.25	73	27.6	0	0
2019/06/30	16.04835	10.25	72.9	29	0	0
2019/09/30	17.48545	10	72.6	29.1	0	0
2019/12/31	19.66327	10	73.2	29.1	0	0
2020/03/31	21.02702	9.416667	73.6	30.1	0	0
2020/06/30	26.17012	7.416667	85.3	23.3	0	0

Source: Author

The sixth tree model for Moody's correctly predicted the classification of all the observations, that is, the predicted class was less stable ("0").

<u>Prediction using SNPoor model Continued</u>

Figure E11 below shows the predicted class for SNPoor using tree two.

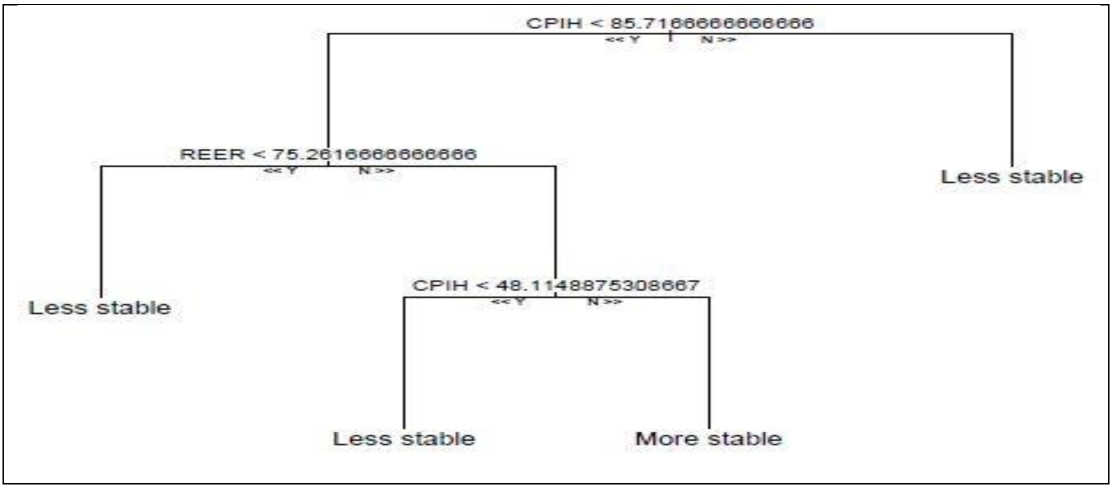


Figure E11: Prediction tree for SNPoor Model Tree 2
Source: By Author

The SNPoor model tree 2 classified the observations using four paths: CPIH and REER.

Path 1: If CPIH is less than 85.7167 and REER is less than 75.2617, the observation is classified as less stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Path 2: If CPIH is less than 85.7167, REER is more than 75.2617, and CPIH is less than 48.1149, the observation is classified as less stable.

Path 3: If CPIH is less than 85.7167, REER is more than 75.2617, and CPIH is more than 48.1149, then the observation is classified as more stable.

Path 4: If CPIH is more than 85.7167, the observation is classified as less stable.

All the observations had CPIH of more than 85.7167 and were classified as less stable. Using the tree generated in Figure E11, the predicted class and the current class are shown in Table EE12.

Table EE12: SNPoor Prediction Model for Tree 2

Time	CPIH	REER	SNPoor Class	Predicted SNPoor Class Tree 2
2019/03/31	110.1	93.57	0	0
2019/06/30	112.0333	92.01333	0	0
2019/09/30	113.1	91.33667	0	0
2019/12/31	113.5667	91.28667	0	0
2020/03/31	114.9667	88.19333	0	0
2020/06/30	114.7333	76.50333	0	0

Source: On Author

The second tree model for SNPoor correctly predicted the classification of all the observations, that is, the predicted class was less stable (“0”).

Figure E12 below shows the predicted class for SNPoor using tree three.

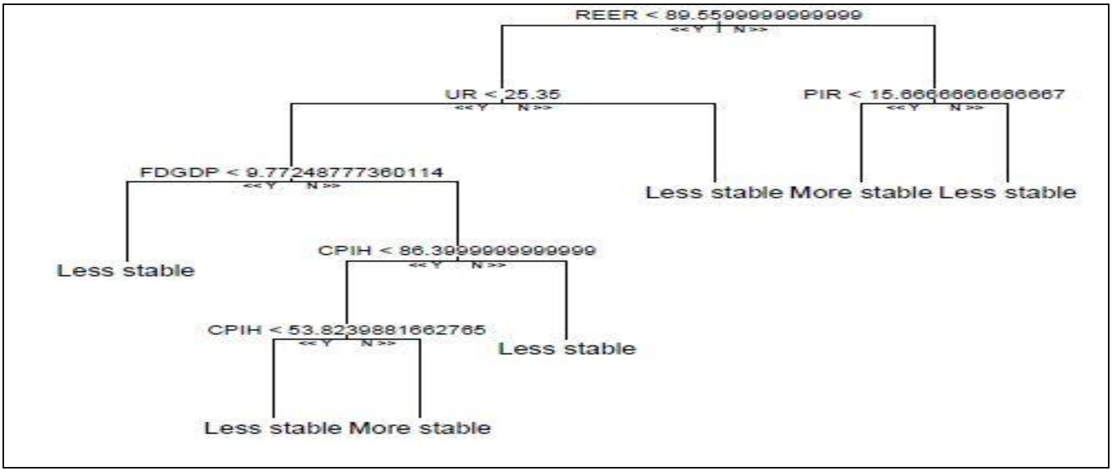


Figure E12: Prediction tree for SNPoor Model Tree 3
Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

The classification of the future observations for SNPoor model tree 3 was done using the variables REER, UR, FDGDP, CPIH, and PIR using seven paths.

Path 1: If REER is less than 89.56, UR is less than 25.35 and FDGDP is less than 9.7725, the observation is less stable.

Path 2: If REER is less than 89.56, UR is less than 25.35, FDGDP is more than 9.7725, CPIH is less than 86.40, and CPIH is less than 53.8240, then the observation is classified as less stable.

Path 3: If REER is less than 89.56, UR is less than 25.35, FDGDP is more than 9.7725, CPIH is less than 86.40, and CPIH is more than 53.8240, then the observation is classified as more stable.

Path 4: If REER is less than 89.56, UR is less than 25.35, FDGDP is more than 9.7725, and CPIH is more than 86.40, the observation is classified as less stable.

Path 5: If REER is less than 89.56 and UR is more than 25.35, the observation is classified as less stable.

Path 6: If REER is more than 89.56 and PIR is less than 15.67, the classification is more stable.

Path 7: If REER is more than 89.56 and PIR is more than 15.67, the classification is less stable.

Observations 1 to 4 had REER more than 89.56 and PIR less than 15.67 and were classified as more stable. Observations 5 had REER less than 89.56 and UR more than 25.35 and was classified as less stable. Observation 6 had REER less than 89.56, UR less than 25.35, FDGDP more than 9.7725, CPIH more than 86.40 and was classified as less stable. Using the tree generated in Figure E12, the predicted class and the current class are shown in Table EE13.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 28-87) **Doi:** 10.34109/ijefs. 202220002

Table EE13: SNPoor Prediction Model for Tree 3

Time	REER	UR	PIR	FDGDP	CPIH	SNPoor Class	Predicted SNPoor Class Tree 3
2019/03/31	93.57	27.6	10.25	17.04585259	110.1	0	1
2019/06/30	92.01333	29	10.25	16.04835113	112.0333333	0	1
2019/09/30	91.33667	29.1	10	17.48544595	113.1	0	1
2019/12/31	91.28667	29.1	10	19.6632713	113.5666667	0	1
2020/03/31	88.19333	30.1	9.416666667	21.02701934	114.9666667	0	0
2020/06/30	76.50333	23.3	7.416666667	26.17011509	114.7333333	0	0

Source: By Author

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 28-87) **Doi:** 10.34109/ijefs. 202220002

The third tree model for SNPoor correctly misclassified the first four observations and was predicted as more stable ("1") instead of less stable ("0"). The determining variable was PIR since it was less than 15, however, PIR was not one of the top four variables that were considered as being important.

Figure E13 below shows the predicted class for SNPoor using tree four.

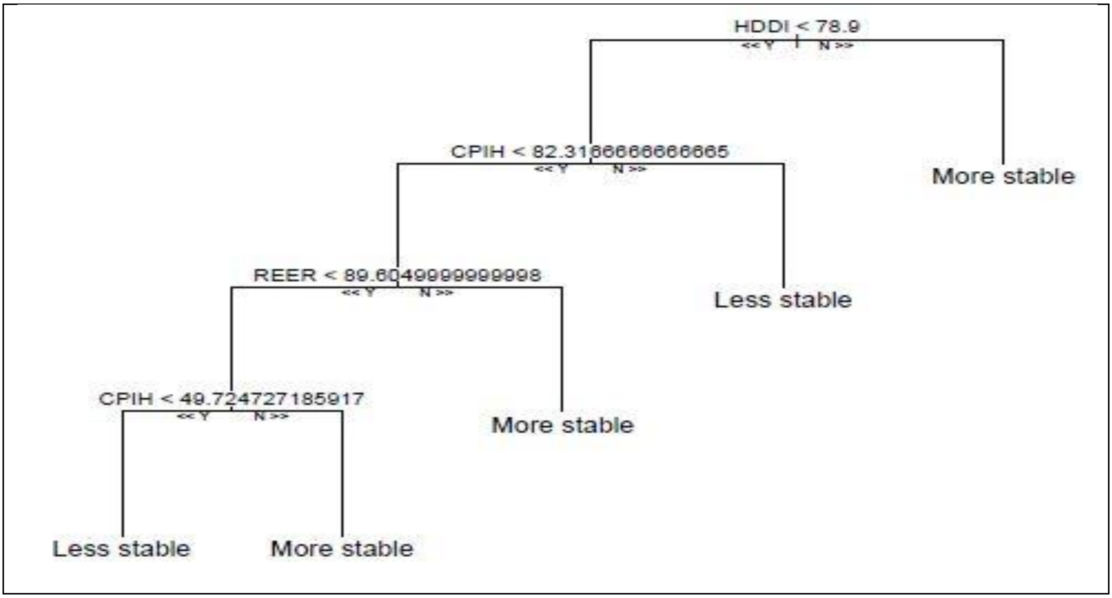


Figure E13: Prediction tree for SNPoor Model Tree 4
Source: By Author

The classification of the future observations for SNPoor model tree 4 was done using the variables HDDI, CPIH, and REER using five paths.

Path 1: If HDDI is less than 78.9, CPIH is less than 82.3167, REER is less than 89.6050, and CPIH is less than 49.7247, the observation is classified as less stable.

Path 2: If HDDI is less than 78.9, CPIH is less than 82.3167, REER is less than 89.6050, and CPIH is more than 49.7247, the observation is classified as more stable.

Path 3: If HDDI is less than 78.9, CPIH is less than 82.3167, and REER is more than 89.6050, then the observation is classified as more stable.

Path 4: If HDDI is less than 78.9 and CPIH is more than 82.3167, the observation is classified as less stable.

Path 5: If HDDI is more than 78.9, the observation is classified as more stable.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Observations 1 to 5 had HDDI less than 78.9 and CPIH more than 82.3167 and were classified as less stable. Observations 6 had HDDI of more than 78.9 and was classified as more stable. Using the tree generated in Figure E13, the predicted class and the current class are shown in Table EE14.

Table EE14: SNPoor Prediction Model for Tree 4

Time	HDDI	CPIH	REER	SNPoor Class	Predicted SNPoor Class Tree 4
2019/03/31	73	110.1	93.57	0	0
2019/06/30	72.9	112.0333	92.01333333	0	0
2019/09/30	72.6	113.1	91.33666667	0	0
2019/12/31	73.2	113.5667	91.28666667	0	0
2020/03/31	73.6	114.9667	88.19333333	0	0
2020/06/30	85.3	114.7333	76.50333333	0	1

Source: On Author

The fourth tree model for SNPoor correctly predicted the classification of five out of the six observations. Only the model classified the last observation as more stable ("1").

Figure E14 below shows the predicted class for SNPoor using tree five.

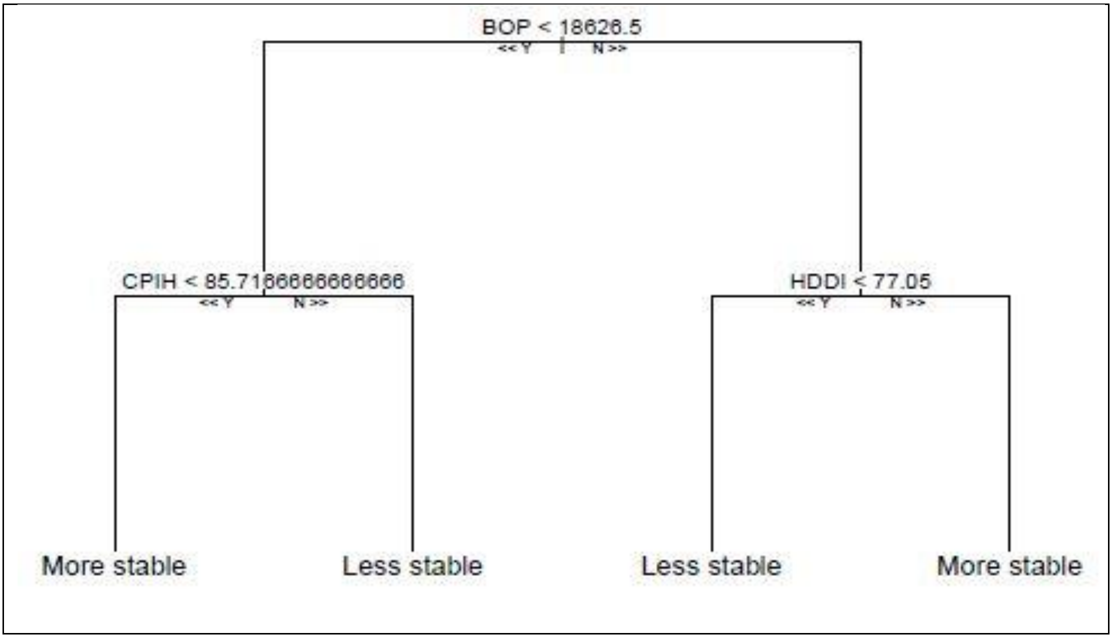


Figure E14: Prediction tree for SNPoor Model Tree 5
Source: By Author

The future observations for SNPoor model tree 5 were classified using the variables BOP, CPIH, and HDDI using four paths.

Path 1: If BOP is less than 18626.5 and CPIH is less than 85.7167, the observation is classified as more stable.

Path 2: If BOP is less than 18626.5 and CPIH is more than 85.7167, the observation is classified as less stable.

Path 3: If BOP is more than 18626.5 and HDDI is less than 77.05, the observation is classified as less stable.

Path 4: If BOP is more than 18626.5 and HDDI is more than 77.05, the observation is classified as more stable.

Observations 1, 3, 4, and 5 had BOP more than 18 626.5 and CPIH more than 85.7167 and were classified as less stable. Observation 2 has BOP less than 18626.5 and CPIH more than 85.7167 and was classified as less stable. Observation 6 had BOP of more than 18 626.5 and HDDI more than 77.05 and was classified as more stable. Using the tree generated in Figure E14, the predicted class and the current class are shown in Table EE15.

Table EE15: SNPoor Prediction Model for Tree 5

Time	BOP	CPIH	HDDI	SNPoor Class	Predicted SNPoor Class Tree 5
2019/03/31	42301	110.1	73	0	0
2019/06/30	-31495	112.0333	72.9	0	0
2019/09/30	43992	113.1	72.6	0	0
2019/12/31	102486	113.5667	73.2	0	0
2020/03/31	201663	114.9667	73.6	0	0
2020/06/30	91508	114.7333	85.3	0	1

Source: By Author

The fifth tree model for SNPoor correctly predicted the classification of five out of the six observations. Only the last observation was classified as more stable by the model (“1”).

Figure E15 below shows the predicted class for SNPoor using tree six.

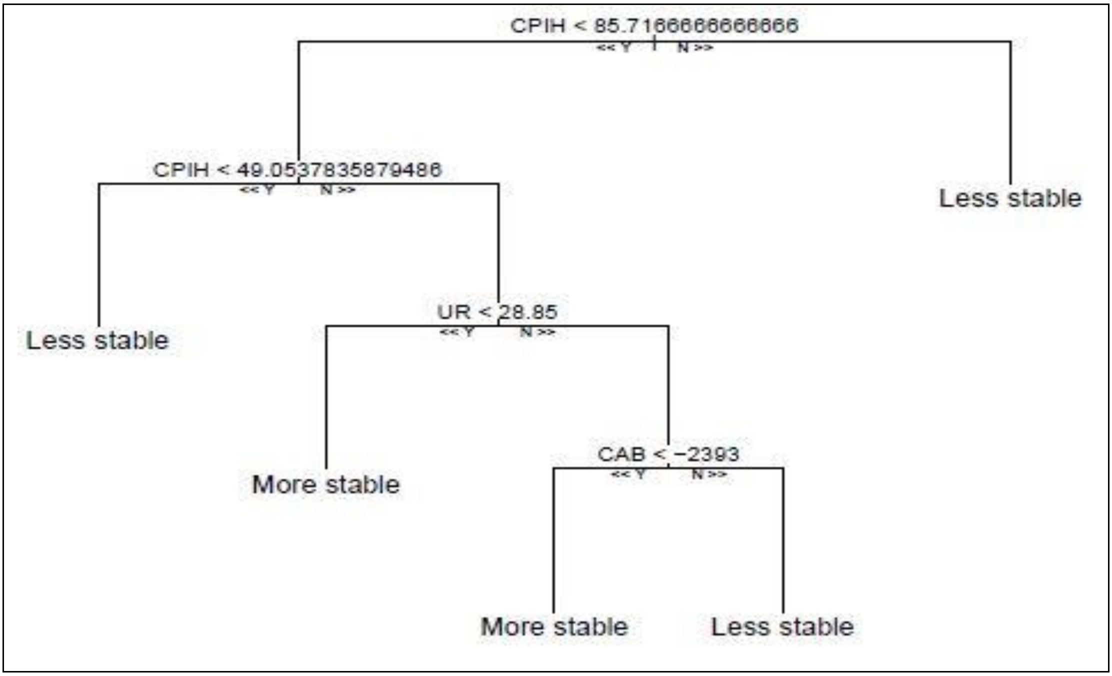


Figure E15: Prediction tree for SNPoor Model Tree 6
Source: By Author

The SNPoor model tree 6 classified the observations using five paths: CPIH, UR, and CAB.

Path 1: If CPIH is less than 85.7167 and CPIH is less than 49.0538, the observation is classified as less stable.

Path 2: If CPIH is less than 85.7167, CPIH is more than 49.0538 and UR is less than 28.85, the observation is classified as more stable.

Path 4: If CPIH is less than 85.7167, CPIH is more than 49.0538, UR is more than 28.85, and CAB is less than -2393, then the observation is classified as more stable.

Path 4: If CPIH is less than 85.7167, CPIH is more than 49.0538, UR is more than 28.85, and CAB is more than -2393, then the observation is classified as less stable.

Path 5: If CPIH is more than 85.7167, the observation is less stable.

All the observations had CPIH of more than 85.7167 and were classified as less stable. Using the tree generated in Figure E15, the predicted class and the current class are shown in Table EE16.

Vol: 14 **No:** 1 **Year:** 2022 **ISSN**: 1309-8055 (Online) (**pp.** 29-87) **Doi:** 10.34109/ijefs. 202220002

Table EE16: SNPoor Prediction Model for Tree 6

Time	CPIH	UR	CAB	SNPoor Class	Predicted SNPoor Class Tree 6
2019/03/31	110.1	27.6	-57063.9368	0	0
2019/06/30	112.0333	29	-29185.8365	0	0
2019/09/30	113.1	29.1	-62178.4028	0	0
2019/12/31	113.5667	29.1	57.3302	0	0
2020/03/31	114.9667	30.1	-8168.3622	0	0
2020/06/30	114.7333	23.3	-8503.0976	0	0

Source: By Author

The sixth tree model for SNPoor correctly predicted the classification of all the observations, that is, the predicted class was less stable (“0”).