International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Decision Tree: Categorizing Financial Inclusion

TanuTiwari, AlpanaSrivastava, Surendra Kumar

Abstract: ***Financial Inclusion (FI) is a global concern and even developed economies are trying to achieve complete inclusion.The inclusion index is reported by many institutions and regulatory bodies considering only one or two key attributes in their reports and hence, the impact of other financial parameters is missed. Further, the reports display an aggregated value at national level. Deciphering the inclusion at individual level will help to take corrective measures and in designing new policies. This study aims to propose a decision ruleusing techniques from data analytics to segment the population into excluded and included. The consolidated weighted scoring method was used over four key financial attributes to identify the actual class.C5.0 algorithm has been applied to arrive at the decision rule which employs technique of entropy or information gain. Surveyed data with 691 records was partitioned into training (80%) and test (20%) data sets. The classification accuracy over the test data set was found to be 100%.The findings of this study could be used by policymakers for individual estimate of FI score and prioritizing the policies.***

Keywords: Financial Inclusion, C5.0, Classification tree, Decision rule

I. INTRODUCTION

Financial Inclusion (FI) in its broader sense means that every individual or entity within the state has an opportunity to access, choose and use all available financial products and services. World Bank defines FI as "Financial inclusion means that individuals and businesses have access to useful and affordable financial products and services that meet their needs – transactions, payments, savings, credit and insurance – delivered in a responsible and sustainable way" (World Bank, 2018). Reserve Bank of India (RBI), the banking regulatory body of India has promoted FI in its terms "Financial inclusion may be defined as the process of ensuring access to financial services and timely and adequate credit where needed by vulnerable groups such as weaker sections and low income groups at an affordable cost**"**(SIDBI, 2008). India, primarily being an agrarian economy(IBEF, 2018)and50**%** of its adults directly or indirectly involved into agriculture or related activities(Sunder, 2018). These peoples have monthly income of **6426 INR** with approximately 94 dollar equivalent. Majority of these people along with other low income strata, do not have access to any formal banking and financial institutions.

Manuscript published on November 30, 2019.
* Correspondence Author
TanuTiwari, pursuing Ph.D. (Management) at Amity Business School
Dr. AlpanaSrivastava, professor at Amity University, Lucknow..
Dr. AlpanaSrivastava**,** professor at Amity University, Lucknow.

Indian government has initiated many schemes to connect these low income strata and downtrodden into main stream of banking and finance. No-Frills account in the name of PradhanMantri Jan DhanYojna (PMJDY) is the most popular and highly accepted among these schemes. PMJDY has managed to open 31.83 crore saving bank accounts within a stretch of 4 years (2014 to 2018) and thus has surpassed 80% bank account penetration from 53% in the year 2014 (Jain, 2018).

Necessity of bank accounts cannot be ignored as they are very helpful in Direct Benefit Transfer (DBT) of funds under various subsidy and social benefit schemes andthus bank accounts are important for every citizen in their own right. However, when it comes to FI, we often adhere to its narrow definition of having bank account rather than its broader aspect of being banked through all kind of financial services like access to debt facilities, being insured as well as being able to use recent innovative digital technologies adopted by banks and promoted as Digital Financial Inclusion (DFI) under Digital India program. Having such a low income, does these people need merely a bank account for making transactions without enough money.

Financial Inclusion within any economy is important to protect the rights of deprived people and facilitate them with the necessary services and their basics needs for livelihood. FI is a major problem globally and apart from under-developed and developing countries, even well developed nations are facing challenges in this line. The 2014 Global Findex report (Demirguc-Kunt, Klapper, Singer, & Oudheusden, 2014) reveals that 38 percent of the world population is still without access to any financial service. Only 27 percent have formal saving and 11 percent have borrowing among those with a typical saving account (Global Partnership for Financial Inclusion, 2018). However, in the period of 2011 to 2014, we have been able to include 700 million people within the banking sphere- a figure to make us happy. But its flipside, 2 billion of the world population is striving without any financial assistance from any agency or institution is enough to make us fateful by marking our under performance in terms of financial inclusion. Uniform offering of financial services is one side of FI, its access at individual level and use is rather restricted for females within the country. Role of women is very vital to achieve the objective of FI and also to improve the national income and growth of economy. This exclusion results into restraining poor's from accessing financial services, mainly loan for their agriculture and business requirements.

When assessing and quantifying FIwe often miss the true picture. We do not adopt a comprehensive approach to measure the FI by incorporating all of its determinants. Considering only bank account as an indicator of financial inclusion is neither correct nor justifiable. We need to adopt a holistic approach in measuring the FI index.

Retrieval Number: D8979118419/2019©BEIESP
DOI:10.35940/ijrte.D8979.118419
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

For this purpose we must give due consideration to major financial instruments mainly bank account, access to loan and debt services, being insured (personal as well as business and property). A recent attribute, access to digital technology for financial operations cannot be missed while supplying any FI index. These variables should be given due importance with properly assigned weights.

The system of financial services helps in utilizing more resources for other fruitful purposes leading to more economic growth, reduction in poverty and other opportunities.(SETHY, 2016). FI can raise labor and firm reinforcement, in turn making government profits boost up and social security strengthening(Aggarwal, 2014). Through financial inclusion government enable social security benefits by transferring subsidies directly to beneficiary's bank account and indirectly reducing the leakages in welfare schemes(Bhatia, 2015).

Financial inclusion is a tool to induce money supply in the economy and to achieve inclusive growth through different schemes launched by government and Reserve bank of India like Kisan Credit card, Business Correspondent, Direct Benefit Transfer and other so as to help poor get into the financial platform or umbrella (Kaur & Abrol, 2018). The important parameter that we need to examine is the measurement of progress of Financial Inclusion programs and how much they are impacting the people who have to be benefitted and also the outreach and results of these schemes(Arora, 2010).

The financial inclusionstudy presumes the exclusion constituents and profiling exercise is important process so as to know the extent of exclusion through usage of index and not by going to each place for measurement(CAFRAL, 2012). Generally financial exclusion takes place when organizations that are expected to incorporate the poor but they fail to achieve it. This study aims to give a true measure of financial inclusion through use of four basic determinants of financial inclusion viz. ownership of a bank account, access to loan requirements, being insures and access to digital banking.

Due to unawareness about financial products existing in the market many households and individuals remain out of financial market purview this condition also called as involuntary non-inclusion. On other hand we have voluntary non-inclusion which means the availability of the financial product in the market but voluntarily individuals don't use it(Banerjee, Kumar, & Philip, 2017). The evolving markets undergo from resource restrictions and informational inadequate market structure. Rather than displaying an aggregated national figure we need to identify those, who are actually excluded from the financial system. Identification at individual level will help in policy formulation and prioritization. In this study it has been tried to identify exclusion using techniques from advance analytics and data science.

II. REVIEW OF LITERATURE

The paper by (Venkataramani,2012), (Honohan, 2005), (Kunt & Klapper, 2011) suggest variables like payments, savings mobilization, monitoring of users of funds for measurement of FI index. However, Beck, Kunt and Peria (2007), MandiraSarma (2008), Rahman, (2013), (Sriram & Sundaram, 2015) proposes variables like access and possibility, availability, penetration, actual use, take up rate, satisfaction, conveniently as variables while other researchers (Wang & Guan, 2017) explains individual's income, education and use of communications equipment, (Ambarkhane, Singh, & Venkataramani, 2016) took population growth, law and order situation and corruption as variables. Then Global Findex (2014) analyzed saves, borrow, make payments, and manage risk, Mehrotra et al. (2009) used figure of rural offices, deposits accounts, and volumes of deposit and credit accounts from banks Chakravarty and Pal (2010) used axiomatic measurement approach for measuring FI.

The paper by Thankom&Rajalaxmi, (2015)incorporates different estimates such as supportability of delivery of services in remote location though defined financial policies by government and banks, also to magnify the financial abilities of financial institutions to enhance financial inclusion.

The paper by Arturo & Frederic ,(1998)infers that financial variables can usefully complement the models and different forecast can serve as reliable and quick checks for making more predictions, The yield curve result obtained suggests that stock process are increasing and other measures play important role in enhancing economic development.

The paper by Hameedu, (2014)has concluded major challenges like the implementation process, models, policies and concepts are still on development stage. Secondly the quantitative analysis through usage of statistical techniques for measurement of progress and enhancement of benchmarking standards are most complex process. Thirdly there is great need for focusing on the distributional and micro dimensions then we also need to explore the business models of current banking institutions so as to get an idea whether they have moved from traditional model to consumer centric models fully or not. These above listed challenges require government and other organizations to expand their area of scope so as to measure the progress.

Chakrabarty (2012) has stated that the ICT (Information and Communication Technology) based Business Correspondent model still remain a challenge and it needs time to stabilize. While financial awareness and literacy still a major topic for concern. This requires coordination among various stakeholders like banks, government organizations, civil society institutions and other. Financial exclusion is an issue among various countries around the globe but each country has to develop its own customized plan of action drawing up from experiences of other countries.

Hannig& Jansen, (2010) concluded that financial innovations by financial organizations and others have made considerable and significant changes in low income segment and also financial goals have been achieved to some extent. But customer protection for those using digital platform has to be taken seriously and regulators are making efforts for these.

III. MOTIVATION

In spite of many government programs andinitiatives, though penetration of saving bank account has increased over the time; the actual financial inclusion has not increased to that level. People are either unable to operate their bank account or they don't need it due to lack of fund.

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

However, every individual with a bank account is categorized as financially included. To differentiate between inclusion and true inclusion, we need a method or rule to classify into included and excluded. It will help the combat to design and implement targeted and customized program to escalate the financial inclusion.

IV. RESEARCH PROBLEM

Studies conducted in this domain mainly measure an aggregated national index for financial inclusion. In this study, it has been tried to classify individuals into included or excluded using predictive modeling techniques and analytics. The problem objective is to formulate a decision rule based on set of demographic, social and economic attribute in order to put them into a particular class (included or excluded).

V. DATA

Using standard questionnaire, primary data has been collected from randomly selected locations from provinces of Uttar Pradesh, India. The questionnaire is based on extensive literature review, expert's opinion as well as recommendations from a focus group to ensure its content validity.The questionnaire was supplied in both Hindi and English to cater linguistic bias. Altogether data was collected from 21 districts accounting 691 responses in all. The data includes demographic attributes (i.e. education, age, income, occupation etc.), social, economic and behavioral attributes.

VI. METHODOLOGY

In this predictive modelingit has been tried to formulate a decision rule, using demographic and socio-economic variables as input and weighted sum of, access to four major attributes of financial inclusion (bank account, digital payment, loan access and insurance) as outcome variable. Weights are obtained as proportion of YES (1) category from each attributes. The weights are normalized so that they range between 0 and 1 which can be interpreted in terms of probability as well.Final score was calculated by multiplying the weights with respective values and summing them. A score of 0.5287 rounded to 0.53 (sum of bank account and credit facility) is chosen as cut off point. Values above the cut off are categorized as included and excluded otherwise. The data set is partitioned into training set (80%) and test set (20%) using simple random method. The decision rule is formulated using the training data set and accuracy of the rule is tested over the test data set. The classification rule is developed using C5.0 algorithm which employs the concept of entropy or information gain for optimal split of the data for growing the tree.

C5.0: It is a classification method to formulate a decision trees and rules which are nearly easy and transparent in there interpretations. C5.0 algorithm produces tree by splitting data into as many branches as the distinct categories. It utilizes the concept of entropy and information gain to measure the node homogeneity as:

$$\text{Entropy (E)} = \sum_{i=1}^{C} -P_i \log_2(P_i)$$

Where, P_i proportion of values falling into class level I and C = number of class levels. After each split, entropy is combined as$\text{Entropy (E)} = \sum_{i=1}^{n} w_i \text{ Entropy}(P_i)$. Finally the information gain is obtained by:

$$\text{Information Gain} = \text{Entropy (Before Split)} - \text{Entropy (After Split)}$$

The algorithm chooses best split which maximizes information gain and this process continues until pure node is obtained or no split is possible.

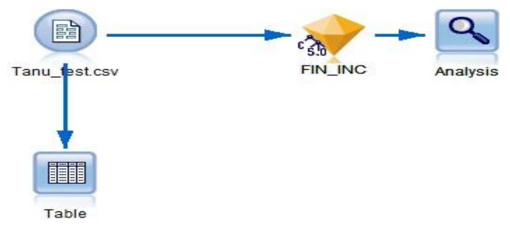
VII. ANALYSIS AND RESULTS

Feeding the data into IBM SPSS modeler and attaching a C5.0 node to it from classification tab results into our decision rule. FIN_INC, our outcome variable was set as output. The modeling not was connected to our test data set and analysis node was attached to judge the predictive accuracy of the model so generated. The model has appreciably wonderful predictive accuracy both over training as well as the test data.

Figure 1: IBM SPSS Modeler C5.0 stream

A) Training Data Set

B) Test data set



VIII. ANALYSIS

The model generated using training data set (see Figure 2) and accompanying decision tree (see Appendix A) is very transparent and easy to understand.

Figure 2: Decision Rule

BANK_ACC = 1 [Mode: YES]

USE_DP = 1 [Mode: YES] => YES

USE_DP = 0 [Mode: YES]

LOAN_ACCESS = 1 [Mode: YES] => YES

LOAN_ACCESS = 0 [Mode: NO]

INSURED = 1 [Mode: YES] => YES

INSURED = 0 [Mode: NO] => NO

BANK_ACC = 0 [Mode: NO] => NO

The analysis node attached to each of the stream (training and test) gives the predictive accuracy of the model generated over the two data sets. It automatically creates the predicted class variable in the data set ($C-FIN_INC) and compare it against original outcome class. On training data the predictive accuracy was found to be 100 percent (**see Table1a**). This was a bit surprising, hinting towards learning idiosyncratic behavior. However, with the test data set, the predictive accuracy is again found to be 100 percent (**see Table 1b)**.

Table: 1a Prediction on Train Data

Correct	555	100%
Wrong	0	0%
Total	555	

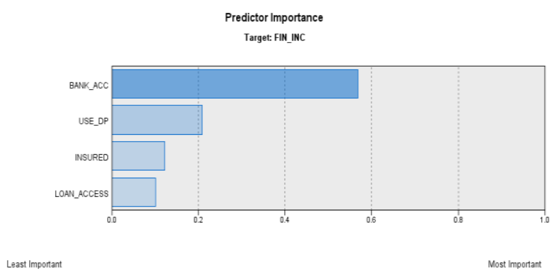
Table: 1b Prediction on Test Data

Correct	136	100%
Wrong	0	0%
Total		136

IX. RESULTS AND DISCUSSION

There are691 responses in all which was divided into training (555, 80%) for modeling purpose and the remaining, into test data set (136, 20%) to check the accuracy of the model (decision rule). Result suggests that the decision rule inducted (**see Figure 2 and Appendix A**) is able to predict the inclusiveness with 100 percent accuracy. The model suggests that bank account ownership is the most important predictor of financial inclusion followed by use of digital payment methods, being insured and access to loan. IBM SPSS modeler also displays a variable importance exhibit (**see Figure 3**).

Figure 3: Variable Importance



X. CONCLUSION

The core agenda of policy makers and regulators has been more focused on making a feasible place for consumers to do financial transactions so that more progress can be insured. Now, having reached to an appreciably good level of inclusion, we need to fine tune the all-inclusive policies. Rather than reporting a gross figure at national level, which fails to account the individual class, now we need a mechanism which can identify and classify the excluded segment from others. For this purpose, a classification method most suitably, a simpler but robust like decision tree is a good option. In this study, the decision rule formulated is showing 100 percent accuracy. This model could be used by policy formulators to distinguish between the two, in order to prioritize the weaker and excluded section. It will help greatly in identifying excluded and targeting for bringing them into financial inclusion through specialized and personalized offerings and schemes.

RECOMMENDATION

The study suggests that indicating the percentage of bank account holders as an indicator of financial inclusion is a vague practice. In order to get a clear and precise measure, we should first identify the basic components of financial inclusion and give due weight to each. Also, rather reporting a national figure, in this data driven world, we need to find an individual score or classification. The current attempt of classification using C5.0 could be adopted by policy makers and financial institutions for policy prioritization and targeted programs for financial inclusion for excluded and deprived people. The study can be extended further by incorporating more sophisticated algorithms over real time data sources from governments' data repositories and warehouses. This study is one of its kind and newer data analytics methods and practices can be adopted for the purpose of personalized identification and offerings.

REFERENCE

1. Ambarkhane, D., Singh, A. S., & Venkataramani, B. (2016). Developing a Comprehensive Financial Inclusion Index. Management and Labour Studies, 41(3), 216-235.
2. Banerjee, A., Kumar, K., & Philip, D. (2017, September). Financial Literacy, Awareness and Inclusion. www.isid.ac.in.
3. Bhatia, S. &. (2015). Financial inclusion–A path to sustainable growth. International Journal of Science Technology & Management, 4.
4. CAFRAL. (2012). Measuring Financial Inclusion from the Demand Side. (pp. 1-9). National Institute of Bank Management.
5. Demirguc-Kunt, A., Klapper, L., Singer, D., & Oudheusden, P. V. (2014). The Global Findex Database 2014: Measuring Financial Inclusion Around the World. Washington, DC: World Bank.
6. Global Partnership for Financial Inclusion. (2018). Argentina's G20 Presidency 2018 priorities paper. Global Partnership for Financial Inclusion.
7. Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2008). Churn Prediction: Does Technology Matter? International Journal of Industrial and Manufacturing Engineering, 2(4), 524-530.
8. Hameedu, M. S. (2014). Financial Inclusion - Issues in Measurement and Analysis. International Journal of Current Research and Academic Reviewl, 116-124.
9. Hannig, A., & Jansen, S. (2010). Financial Inclusion and Financial Stability: Current Policy Issues. ADBI Working Paper Series.
10. IBEF. (2018, June). Agriculture in India: Information About Indian Agriculture & Its Importance. Retrieved July 2018, from www.ibef.org: https://www.ibef.org/industry/agriculture-india.aspx
11. Kunt, A. D., & Klapper, L. (2011). Measuring Financial Inclusion. The Global Financial Inclusion Index (Global Findex).
12. Saha, D. (2017, June). 70% Of India's Farm Families Spend More Than They Earn–Debt Main Cause of Suicides. Retrieved July 2018, from http://www.indiaspend.com: http://www.indiaspend.com/cover-story/70-of-indias-farm-families-spend-more-than-they-earn-debt-main-cause-of-suicides-26738
13. SIDBI. (2008). The Committee on Financial Inclusion. https://www.sidbi.in/files/Rangarajan-Commitee-report-on-Financial-Inclusion.pdf.
14. https://www.financialexpress.com/budget/india-economic-survey-2018-for-farmers-agriculture-gdp-msp/1034266/
15. Thankom , A., & Rajalaxmi , K. (2015). Financial inclusion: Policies and practices. IIMB Management Review, 267-287.
16. TRAI. (2018). Highlights of Telecom Subscription Data as on 31st December, 2017. NEW DELHI: Telecom Regaulatory Authority of India.

17. Venkataramani, B., Gupte, R., & Gupta, D. (2012). Computation of financial inclusion index for India. Procedia-Social and Behavioral Sciences, 133-149.
18. World Bank. (2018 , April 20). Financial Inclusion. Retrieved July 1, 2018, from www.worldbank.org: http://www.worldbank.org/en/topic/financialinclusion/overview

AUTHORS PROFILE

TanuTiwari is B.Com and MBA(Finance), currently pursuing Ph.D. (Management) at Amity Business School. Her research area comprises Financial Inclusion and inequality.
Email- tanutiwari.88@gmail.com

Dr. AlpanaSrivastava, professor at Amity University, Lucknow. She is involved in teaching subjects like economics, econometrics, statistics and operations research at graduation, post-graduation and Ph.D. level.

DrSurendra Kumar is a professor in BBD University, Lucknow. His keen interest is in marketing and operations.

Appendix A