An Analytical Model for Evaluating Social Security Schemes-A Focus on "Ayushman Bharat" Universal Health Scheme in India

P. Sunil Kumar, Raghunatha Sarma, Satya Sai Mudigonda

Abstract*: The government initiated social security schemes in countries such as India, target a large proportion of the population to provide various types of benefits that involve a number of stakeholders. Such schemes are executed by a large number of transactions between the Government agencies and the other stakeholders on a real time basis, thus resulting in large data sets. Current research advancements in the domain of social security schemes include analysis of sequential activities and debt occurrences for such transactions at the national level only. It has been a challenge in recent times to monitor and evaluate the performance of such gigantic schemes which also involves financial decision making at different levels. This paper proposes an innovative frame-work that combines data mining strategies with actuarial techniques to evaluate one of the popular schemes in India, namely AB-PMJAY ("Ayushman Bharat–Pradhan Mantri Jan Arogya Yojana") launched by the Government in 2018 at family level. In the proposed framework, the scheme has been divided into a number of sub-processes for which various data mining techniques such as, clustering, classification, anomaly detection and actuarial techniques for pricing are proposed to evaluate the scheme effective at micro level.

Keywords: classification, clustering, anomaly detection, social security scheme, actuarial techniques, Ayushman Bharat, AB-NHPM.

I. INTRODUCTION

Social security schemes are the provisions and funding of various benefits by the Government for its citizens. The major facilities provided by the Government include access to health care, income security, unemployment, work injury, maternity care, etc. International Labor Office (ILO), Geneva in its report 2017-2019 stresses the need for social protection systems to achieve sustainable development goals for the entire world population [1]. It has also estimated that 71 percent of the current world population is yet to access the benefits of social protection systems. Hence as a part of its agenda-2030, ILO has plans to implement social protection systems for all, including floors, for poverty reduction and prevention.

The agenda also includes universal health coverage as one of the sustainable development goals. As far as universal health coverage is concerned, the report estimates that 56 percent of the world's rural population and about 29 percent of the world's urban population are not covered under any health coverage system [1]. It also states the requirement of 10.3 million global health workers for long term care (LTC), of which 7 million are required for rural and 3 million for urban areas. To fill this gap the ILO, in its agenda-2030 has plans to employ more than 13 million for LTC work.

In India, after post-independence, such schemes were predominantly oriented towards people in organized sectors only. It is important to mentioning here that according to Labor and Employment Department, Government of India, 90 percent of the labor force in India belongs to unorganized sector who contribute more than 50 percent of national GDP [2]. So keeping these facts in view the first social scheme was introduced by the Government of India, for the informal sector in 1995 under the name National Security Assistance Program (NSAP).In due course, every State Government also introduced many schemes in different modes e.g. Insurance mode and Trust mode with partial contribution by the beneficiaries. Charan Singh et. al. enforced the need for social schemes with least or no contribution by the beneficiaries[3]. The National Health Agency (NHA) studies reveal that the major expenditure incurred towards health care is about 68 percent of their monthly income in the unorganized sector [4]. Keeping population size, economic status and out of pocket expenses towards minimal health that exists in the unorganized sector, an immediate need arose to initiate a universal scheme.

One such scheme of National Health System is "Ayushman Bharat–Pradhan Mantri Jan Arogya Yojana" (AB-PMJAY). The purpose of this paper is to propose an analytical model for evaluating this scheme. This assessment model can be extended to any other social security scheme with appropriate modifications.

In section II, we briefly present the research work done in social security domain (SSD).

Revised Manuscript Received on September 25, 2019.

P.Sunil Kumar*, "Mathematics and Computer Science,Sri Sathya Sai Institute of Higher Learning,Muddennahalli, Chickballapur, India. Email: psunilkumar@sssihl.edu.in"

Dr Raghunatha Sarma,"Mathematics and Computer Science, Sri Sathya Sai Institute of Higher Learning, Puttaparthi, India. Email: rraghunathsarma@sssihl.edu.in"

Sathya Sai Mudigonda, "Mathematics and Computer Science,Sri Sathya Sai Institute of Higher Learning,Puttaparthi,India, mail:sathyasaibabamudigonda@sssihl.edu.in"

Section III is about literature review on social security using data mining. Section IV briefs about the motivation to take up this research work. Section V, VI and VII depict AB-PMJAY, its challenges, and opportunities. In Section VIII the analytical model for the AB-PMJAY scheme has been explained.

We conclude with section IX by suggesting some future work that can be done to logically extend the model effectively to various aspects.

II. RESEARCH HISTORY IN SSD

The research work done in social security domain (SSD) can be explored category-wise as well as research approach-wise. Mario and Evangelos in the social security research center, University of Malaya, analyzed 1005 articles published during 1967-2017[5]. They observed that 25 categories of research areas have been explored in SSD using more than 20 research approaches. In this timeline, the early techniques were econometric and statistical only. They also stated the possibility of research work in social security in the domains of health care, poverty control, insurance for pensioners, unemployment, transfers of taxation. The study explored the socio-economic problems that have a direct or indirect impact on the living standards of the people belonging to different geographical areas [5]. However, the report done by S Seng regarding the common challenges in implementing these schemes includes access, coverage, sustainability, adequacy, equitable – redistribution, and social pooling [6].

Mario and Evangelos, developed a comprehensive tool based on economic modeling for evaluating the pension schemes of the countries namely Singapore, Thailand, Malaysia, Indonesia, and the Philippines [7]. This tool classified the pension schemes as high performance, good performance, low performance, and poor performance. This tool was designed on the basis of different formulas, designed using econometric methods.

Denis Garand et al. reviewed the research and publications done by Government bodies on health insurance schemes at Ghana, Thailand, Rwanda, Indonesia and India and proposed a performance evaluation framework based on key performance indicators(KPI's) that are found common in all the schemes [8].

Subramaniam Iyer developed an analytic model for evaluating social security pension schemes using stochastic actuarial evaluation technique [9]. This model used descriptive statistics for the evaluation process.

In recent times the availability of huge data sets has encouraged the use of data mining and machine learning techniques for analysis. These techniques are used to build the models that can learn from the archive data. Even though social security systems provide a huge opportunity for data analysis, according to Lingbong Cao seldom research activity has been achieved in this evolving area [10]. The proposed work in this paper is a progressive step towards the application of data mining techniques and actuarial techniques for social security domain.

III. LITERATURE REVIEW -SSDM

Data mining techniques are used to draw inferences from the large data sets that are being generated at different levels from different sources.

Longbing Cao suggested a framework i.e. Social security Data Mining (SSDM) for the Australian Commonwealth Department of Human Services (Centrelink) that generates huge amounts of data on a daily basis [10]. The framework includes source data transformation, model building, and data mining techniques for model evaluation.

Yanchang Zhao et al. Proposed different applications of data mining in social security domain [11]. Demographic patterns of customers were identified using decision tree and association rules mining. Sequence mining is used to find activity sequence patterns related to debt occurrence. Slow payers and quick payer's patterns are discovered using combined association rules. Clustering and ANOVA are employed to check the effectiveness of a new policy.

Since SSD involves a large population, it is a challenge to identify similarity among customers or beneficiaries. Camila Maione et al. used social data of farmers living in North-Eastern Brazil to estimate natural groups and then analyze their profiles based on certain social criteria namely demographic, economic, agricultural production and food insecurity data [12]. They have used clustering and variable selection techniques for the purpose.

Leitao Zhang et al. used XGBOOST algorithm to predict the subjective well-being of individuals. Social security feature was found to be contributing more towards subjective well-being of the individuals [13]. A comparative analysis was also carried out in with other benchmark algorithms, like K-Nearest Neighbors and Support Vector Machine [13]. The performance of XGBOOST was found better than other algorithms. The information flow in social security schemes is generally carried out in a top to bottom fashion among various institutions of service in the Government. At the national level, the modalities of the schemes are framed and were passed on to state-level bodies for dissemination and execution. The state-level institutions disseminate the scheme related information by using the personnel, at district levels, who are responsible for identifying and registering the beneficiaries, provide awareness and fulfill the services as demanded by the scheme. The stakeholders in these schemes are eligible insurance companies and trusts, who are selected by the institutions at state as well as national level. Since such schemes generate large volume data which is being stored and disseminated, a number of research works are carried out at institutional levels using data mining techniques.

Wafa and Abdallah used algorithms like K-means and SOM (Self-organizing Maps) with K-means to understand customer segmentation in the health insurance [14]. The TIC-2000, CRM data set was taken for the purpose which has 86 nominal attributes and the target variable was binominal. The data set was first tested for favoring clustering tendency by finding the distance between each pair of features.

Finally, Davies Bouldin index was used for comparing both the algorithms.SOM (Self-organizing Maps) with K-means performed better for all the clusters to identify the customer groups.

Seyed Mohammad et al. developed a procedure to measure customer loyalty for a marketing company SAPCO in Iran [15].The procedure used weight and non-weight based K-means algorithms to make clusters based on four parameters namely, recency, monetary frequency, and loyalty. It measured Customer Loyalty Value (CLV) from the resultant clusters. Yang Xie et al. proposed a temporal data mining approach using Bagged Regression trees, to predict the number of days in the hospital-based on claims made by the customers [16]. The categorical and numeric data were converted into time-series data and quarterly bins and yearly bins were used for prediction. The quarterly model performed better than the yearly model. Hyunjung Shin et al. used two measurements namely CDA (composite degree of anomaly) and GDA (Grade for degree of anomaly) for scoring and segmentation of abusive health insurance providers [17]. The model used radial diagrams for scoring which was based on 5 key indicators of abusive billing patterns. Decision trees were used for segmenting the providers based on the CDA scores.

VipulaRawte and Anuradha developed a hybrid approach using ECM (Evolving Clustering Method) and SVM (Support Vector Machine) to detect fraudulent claims in health insurance [18]. The ECM method generates disease-wise clusters of claims. These clusters are then classified into fraudulent or legitimate claims using support vector machines.

IV. MOTIVATION

The Social security schemes launched by the Government of India, involve a number of stakeholders in a hierarchy to provide services to the common citizen. Such schemes are carried out with huge budgetary plans at national/state level covering a large portion of the population. It also performs billions of transactions between the beneficiary and the stakeholders, generating huge amounts of data. In order to analyze the performance of such schemes and to understand the hidden information e.g. beneficiary behavior, segments of progress and financial projections, techniques of data mining and machine learning integrated with actuarial techniques are most suitable. For the proposed work, we have focused on AB-PMJAY, the world's largest National Health Scheme that was launched in India in 2018.

V. OBJECTIVES OF AB-PMJAY

"Ayushman Bharat–Pradhan Mantri Jan Arogya Yojana" (PM-JAY) intends to move the nation closer to 'Universal Health Coverage (UHC) to achieve Sustainable Development Goals (SDG) by providing a health coverage of Rs.5 lakh per family per year with a minimal premium, for secondary and tertiary care hospitalization. It plans to cover 10.74 crore poor and deprived rural families with special concern for girl child and senior citizens (covering approximately 50 crore beneficiaries), who can avail free treatment at public and impaneled private hospitals in a cashless and paperless manner. The scheme does not enforce a cap on family size and age of members, however for initial identification the families should either feature as targeted groups as per the latest "Socio-Economic Caste Census" (SECC) 2011 data (8.03 crore families in rural and 2.33 crore in urban areas)or enrolled under "Rashtriya Swasthya Bima Yojana" (RSBY).

All the existing state health schemes will be subsumed with Ayushman Bharat Scheme. The targeted families will get the benefit of national portability and can also avail services across India. The scheme includes 1,354 medical packages covering surgery, medical and daycare treatments, cost of medicines and diagnostics along with state-wide pre-existing diseases.

• The scheme is to be jointly implemented by central and state governments with a funding pattern of 60:40 respectively, through national and state health agencies (NHA, SHA) in any of the three modes e.g. Insurance mode/Trust mode/Mixed mode adhering to "standardized treatment guidelines" (STGs) and standardized package rates which promises to build a holistic healthcare ecosystem [19].

VI. CURRENT OPPORTUNITIES

Bringing Hospitals, Third Party Administrators (TPA) and Health Insurance Companies on one single platform where data sharing becomes easy and secured among them. This helps avoid fraud due to fake hospitals, medical practitioners and pharmacies. Ultimately an authentic repository of this data can be created for validation purposes.

Data repository that is created during the process of implementation of this scheme will act as a source for various data analytics which in turn helps stakeholders to make appropriate informed decisions.

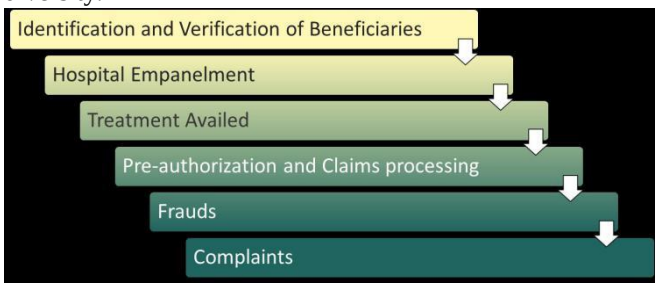
Unique Ids for these services will prevent fraud and increase claim settlement efficiencies.

Scheme created various employment opportunities in direct and indirect ways to main Health Systems available in India. It is expected to create about 100,000 new jobs [19].

VII. IMPLEMENTATION CHALLENGES

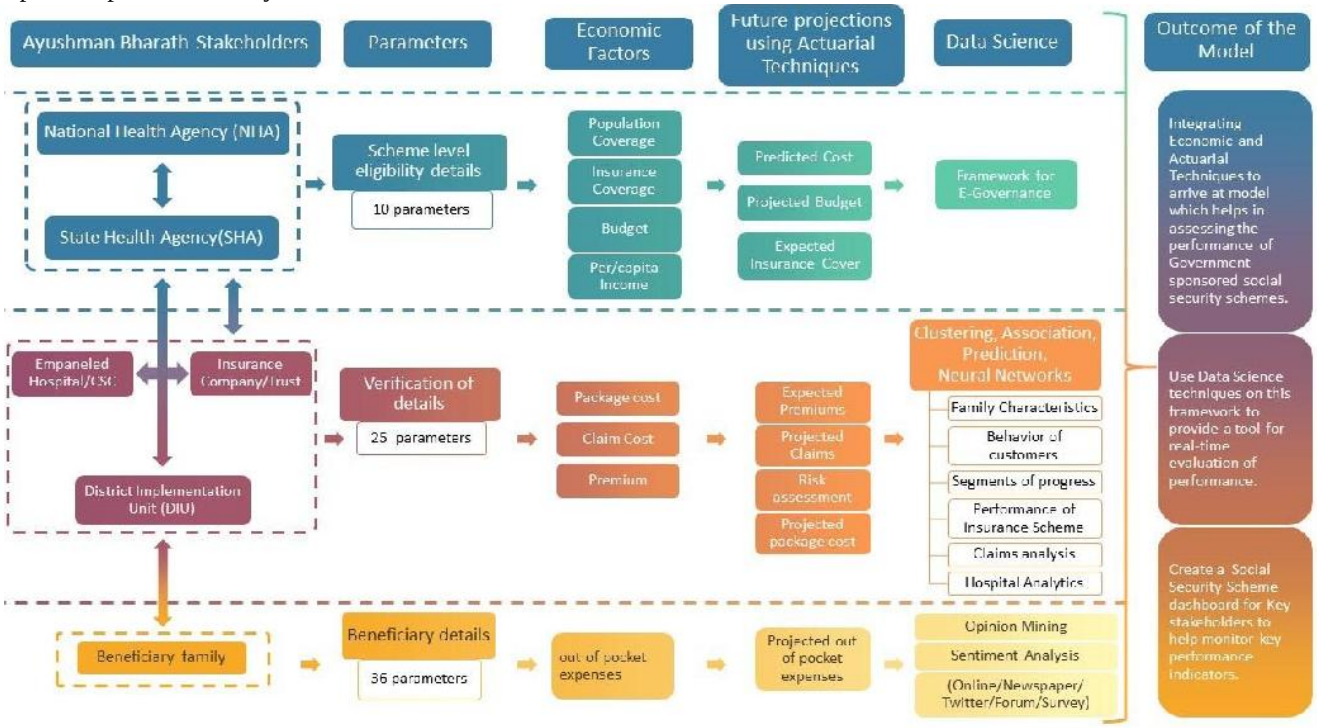
Due to the volume of data and integration of various stakeholder's systems, it may lead to data inconsistencies and technical integration issues.

Defining a suitable process flow that is applicable to all stakeholders in all States and implementation of the same may become a challenge due to technical literacy and language diversity.



(Fig: 1)

Preventing and detecting fraud may become a challenge as information may not be available real-time across the country. Price differences between market price and AB-PMJAY for expensive procedures may



(Fig: 2)

lead to increase funding for the Government and lower satisfaction among beneficiaries. Issues such as duplication of beneficiaries, inefficiencies in scheme implementation, frauds and poor targeting may increase the financial cost of the scheme

VIII. ANALYTICAL APPROACH FOR EVALUATION

For the purpose of assessing the performance of AB-PMJAY, we propose a model in which various stakeholders, parameters, economic factors, actuarial and Data Science techniques are incorporated.

Figure: 2 depicts Analytical Approach for Evaluation

Each part of the assessment model is explained in detail in the following sections.

A) Stakeholders of AB_PMY

National Health Agency (NHA): A body of the Central Government of India that frames the rules and regulations to rule out the scheme AB-PMJAY .It also plays an important role as the only approval authority at various stages during the implementation of the scheme.

State Health Agency (SHA): A body of the State Government that is responsible for integrating the existing state health scheme with AB-PMJAY. Selection of the Insurance companies, Ayushman Mitra (AM), Modes of implementation, empaneling hospitals and other modalities are taken care of by SHA with due approval by NHA.

Empaneled Hospitals: These are the enlisted hospitals (Government /Private) by SHA that provide health services to the beneficiaries.

Insurance Company or Trust: The body that is responsible for approving or rejecting the claim made by the service provider or insurer. District Implementation Unit: The body that is responsible for disseminating the information regarding the scheme and its benefits to the people and then guide them to enroll in the scheme. Beneficiary family: The households of urban or rural India whose details are available in SECC data or have RSBY URN or registered under any state-level health scheme

B) PARAMETERS

The list parameters identified from the official website of AB-PMJAY are as follows:

Beneficiary Detail:

Parameters	Type
ID Proof	Adhar Card, Ration Card, PM Letter-No, RSBY-URN, State-ID,AHL_HH_ID None
Locality-population	Urban, Rural
AB-NHPM Id	Yes, No
Age	Range
Monthly Income	Numeric , range
Mobile Number	Yes, No
Eligibility	SECC, RSBY, State-based-id
Occupation	Daily-wage, Private, Business, Farmer
Deprivation Category	Rural(D1-D5),Urban(O1-O11)
Confidence Score	0-99,Numeric
Awareness Mode	Govt-Person, Media, other
Trust/Insurance Feedback	Selected, Rejected
Exclusion-Scheme	(E1-E14)
Enrolment location	Hospital, Service center, Other

ECHP Verification Details:

Procedure code	Category
Specialty care	Primary, Simple secondary Complex secondary, Tertiary, Emergency
package	Medical, Surgical
pre-authorization	Yes No
Status of enrolment	State specific
Infrastructure availability	Yes ,No
Biometric check up	Success ,Failure
Specialty code	S1 to S16 and M1 to M8
Procedure name	Category
Pre-op Investigation	Text
Pre-op Investigation	Text
Referral status	Taluk, PHI, Dist-PHI, Private-Hospital
Category of treatment	2A,2B,3A,4A
No of Days of treatment	Range
Claim period	1-15 days
Multiple surgical treatments	Yes NO
Term renewal of Hospital	Yes No
Package rate category	medical/ day care/surgical procedures/ interventions
Type of benefit	Hospitalization expense /Day care treatment benefits /Pre and post hospitalization expense /new-born/child care benefit
Rates	(Numeric, Ranges) aggregated data
Insurance Premium	Real number
Over charges	Yes No
Sum assured availability	Yes No
Incidental Investigation	Yes No
Patient-hospitalized	Yes No

Scheme Level Details:

Mode of treatment	Trust, Insurance, Mixed
Region	N,S,E W
Awareness strategy	Person, Media, other
Hospital	Private(profit), Private(Non-profit), Govt.Hospital
Enrolment point	PHI, Enrolment centers, Private Hospital
SHA/ISA verification	Yes, No
AM Available	Yes, No
Enrolment staff availability	Yes, No
Other Insurance benefit	Yes, No

C) Factors specific to Country and State

Country Specific:

- Penetration of insurance
- % GDP spent on health
- Demographic data
- Disease prevalence
- Literacy & Poverty
- Availability of medical facilities and infrastructure
- Culture based lifestyle
- Adoption of alternative medicine

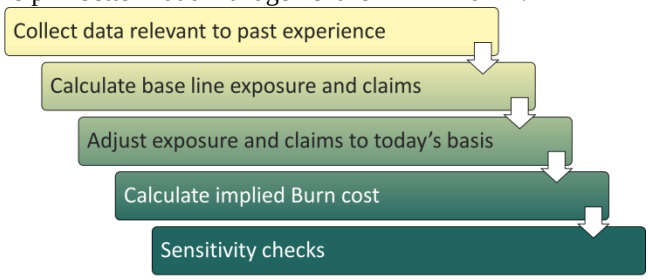
State Specific:

While the coverage being provided by all the states may be the same, the experience might differ due to different factors. When using one state's experience to price another state's scheme, it is important to consider some factors that differentiate the target population in the two states. They are

- Past history of public health care availability. This will help to determine the extent of the awareness of the people to claim.
- Economic status of the people in the state.
- Historical medical scheme's experience. This is a proxy for the general health of the population in the state
- State specific scheme design. For example allowance for transportation provided by the state along with the vase product
- Availability of hospitals that are empaneled under the scheme. This will determine the accessibility hence the frequency for claiming.

D) Actuarial Pricing Methodology for AB-PMJAY

Pricing is one of the most important and challenging actuarial activity for AB-PMJAY. For the prevailing scheme we should first arrive at the best estimate claims cost based on the data available. However, there might be many adjustments that would be required. One of the major adjustments that need to be accounted for would be the differences in the availability of the product design, hospitals, target population, package rates, and package numbers. Fraud detection for these types of schemes are various important as it may act as a major deterrent if left uncontrolled. The integration of actuarial techniques and technology in identifying fraud can help in better fraud management for AB-PMJAY.



(Fig:3)

Factors for Burn cost calculation

- Base claims cost
- Calculate claims experience by period
- Analysis of claims by exposure period to check uniformity by duration
- Allowance for earned premium
- Allowance for IBNR and outstanding claims
- Sensitivity checks with premium of last year and medical indemnity products of similar Sum Insured level

Private vs Public hospitals

- Relativity of claims experience between public and private hospitals using claims severity

- Sense checking and validating relativity with other medical products data
- Allow for changes in the mix of public and private hospitals

District

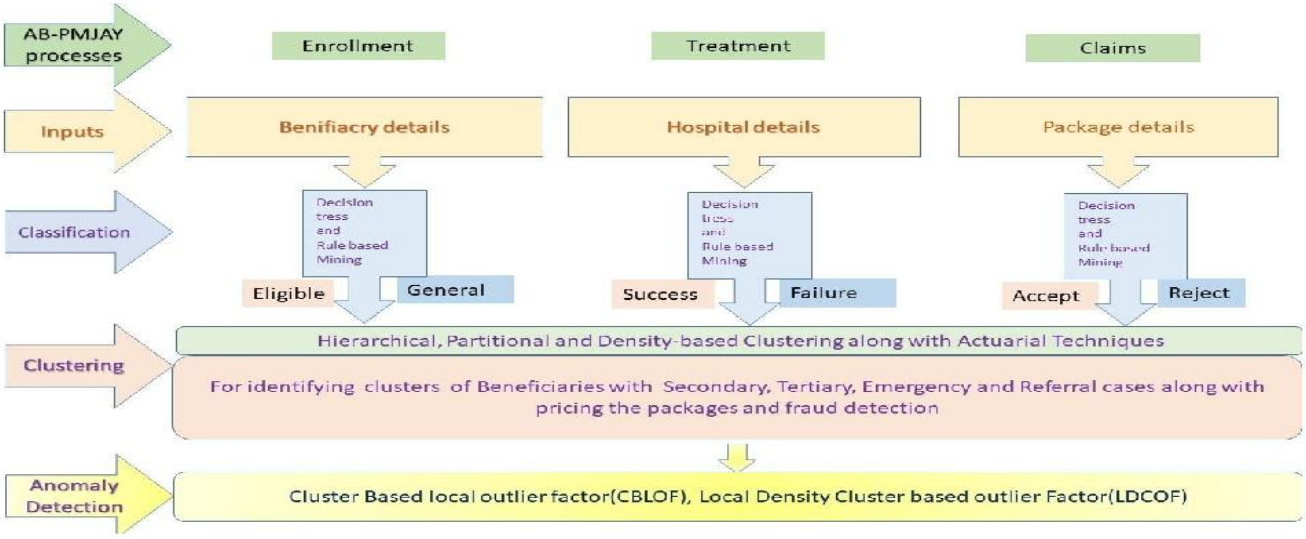
- Relativity of claims experience by district
- Sense check with burn cost relativity and severity relativity from other medical data
- Allow for changes in business mix by district
- Allow for higher utilization and higher severity in cities as compared to underdeveloped districts.

Category of families

- Burn cost relativities for different categories of families.
- Allow for changes in the mix of categories of families

Adjustment for claims repudiation

- Assume lower or higher repudiation of claims based on insurer
- Involvement of external TPA can attract a lower claims repudiation compared to internal TPAs.
- Allow for reduction in fraud due to any risk mitigation measures



(Fig:4)

Allowance for changes in package rates

- Key changes in package rates by section
- Section % allocation of burn cost from other available sources of data

Indications from other data sources

- Any other public scheme experience in Punjab
- Indications from other RSBY and Ayushman Bharat scheme data
- Retail Indemnity medical data

E) Data Science:

Methodology for Data Analysis

Customized Data mining for Ayushman Bharat Scheme:

Various stages in this research endeavor are as follows:

1. Acquisition of data
2. Pre-processing the collected data
3. Building the implementation model
4. Testing the Model

Stage 1: Data acquisition

Collect the data from Government Portals e.g.

- https://secc.gov.in/welcome :

Aggregated information on the said parameters mentioned above is available in Excel format for analysis.

- https://data.gov.in:

Hospital statistics in India staring with PHC to District hospitals are available in CSV format.

- https://pdsportal.nic.in

Information about the "Antyodaya **(AAY)"** ration cards which are issued to" Poorest of poor" households is available in this portal in excel format.

- www.pmjay.gov.in official site of AB-PMJAY:
- The vital statistics regarding package rates and state-wise coverage along with the following data is available for each state which can be utilized for analytics purposes.
- The coverage of state-wide households/families under AB-PMJAY involves the parameters e.g.
- "Total no household, Household Selected under AB-NHPM, rural population, urban population, Automatic Inclusion, Funding pattern, Scheme, Year of launch, Target Population and No of Families covered".
- State-wise health insurance schemes and the parameters involved are "State Scheme, No of beneficiary families, Beneficiary type, Insurance coverage, Premium charged/family, percentage of expenditure on health, No of hospital impaneled, No of packages/procedures, Technology-usage, No of claims, Current implementation mode, Implementation mode under AB-PMJAY".

- Aggregated data related to "number of households, land details, salary/income details, age/gender/caste wise population, income resources, and marital status, educational

Details and disability details" are available at Gram panchayat level from SECC-2011 survey.

- Collect data through questionnaires: From the patients and Nodal officers in district headquarter hospitals.

Stage 2: Pre-processing the collected data

- The dataset contains demographic, aggregated, time series, spatial and transactional data. Since the data sources are collected from multiple sources and secondary in nature, data integration can be done after cleaning and transforming it into a suitable format.
- For this, techniques like feature encoding, standardization, binarization, and normalization can be used to transform the data for analysis purposes.

Stage 3: Building the implementation model

- AB-PMJAY scheme has been divided into three sub-processes in this model. Accordingly in Fig: 4 the scheme has been analyzed at three major levels viz. enrolment, treatment, and claims based on data available for beneficiaries enrolled hospitals and package details respectively.
- Following the rules stated by National Health Agency (NHA) for implementing AB-PMJAY scheme [4], data mining techniques such as decision trees, rule-based mining along with logistic regression will be used to build sequential patterns at all the levels.
- In the first instance, these patterns will classify the beneficiaries as eligible or general. Secondly, for eligible beneficiaries, the activity sequences followed by an Empaneled Health care provider (ECHP) will be classified as successful or unsuccessful. Thirdly, the package cost claimed by the ECHP or the insurer for the beneficiaries and the hospitals identified in previous steps will be classified as accepted or rejected.
- The classified outputs will be grouped into several clusters to analyze the performance of the scheme.
- If the activity sequence was found successful i.e. the beneficiary is eligible, treatment is successful and the claim was accepted then clustering techniques like K-Means, hierarchical and DBSCAN, can be used on the time-series data to group the beneficiaries of a specific region availing different kinds of benefits viz. secondary, tertiary, emergency, pre-authorization and referrals.
- This can be done using the outputs obtained from the classification task as well as on the availability of declaration statements made by the insurer or beneficiary regarding the claims.
- This can be further extended to estimate the efficiency of all the services provided by the scheme by measuring the distance and density within the clusters.
- The clusters thus identified can be used to fix the cost of various packages as needed at different locations. Regression and Actuarial techniques will be applied to these clusters to predict the future cost of premiums, packages, and budgetary issues.
- For some of the clusters, anomaly detection will be carried out based on clustering techniques. The outliers can be utilized for fraud detection or inclusion of more beneficiaries into the scheme.

Stage 4: Testing the Model

- Choosing the best approach model following 20/80 principle
- Calculating coverage and accuracy ratio
- Using the knowledge base of the scheme to establish the model.

The outcome of the model is

- Integrating Economic and Actuarial Techniques to arrive at a model which helps in assessing the performance of Government-sponsored social security schemes.
- Use Data Science techniques on this framework to provide a tool for real-time evaluation of performance.
- Create a Social Security Scheme dashboard for Key stakeholders to help monitor key performance indicators
- To provide cost-benefit analysis for new hospitals (private) enrolled into the scheme.
- Package rate adjustment based on analysis outputs.

F) Technology suitable for managing AB-PMJAY schemes

Storage:

Because of the recent initiatives of Government of India in information technology viz. big data in Governance, Cloud adoption, Request to technology giants to have data centers in India, the amount of information to be stored will increase rapidly. These huge databases which contain information about the beneficiary families, Stockholders contain a wealth of data constitute a potential goldmine of valuable business information. In addition, new applications such as dynamic financial analysis and catastrophe modeling require the storage, retrieval, and analysis of complex multimedia objects, which are often represented by high-dimensional feature vectors.

IX. CONCLUSION AND FUTURE STEPS

The size and scale of AB-PMJAY scheme necessitates an analytical model for assessing its performance on a real-time basis. This helps the Government and other stakeholders involved in implementation to fine-tune their operational efficiency. We believe that the innovative approach presented here will meet this requirement and provide insights to administrators to make appropriate decisions with regard to operational efficiency and improvement to the scheme.

Integrating actuarial techniques with data mining will help monitoring the scheme effectively. As a next step, we suggest a quarterly valuation of AB-PMJAY scheme using the proposed model in the paper and monitor the progress using the analytics provided. Further, we suggest integrating actuarial techniques to predict the future viability of the scheme by region-wise/ state-wise / district-wise for the whole country.

Further analysis of hospitals, disease, drugs, claims, premiums, and fraud can be done using this model based on the availability of data.

ACKNOWLEDGMENT

We express our deep sense of gratitude to the founder chancellor of our university, Bhagawan Sri Sathya Sai Baba for His inspiration and message for life which becomes the ultimate purpose of this proposed work.

REFERENCES

1. "https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_604882.pdf,"
2. International Labor Organization, 2018. [Online]. "https://labour.gov.in/sites/default/files/Report%20vol%204%20final.pdf,"
3. Labor and Emplyement Department,India, 2014. [Online]. S. B. Charan Singh, "Social Security Schemes:
4. A Case for Universalisation," IIMB-WP No 498, 2015. "https://www.pmjay.gov.in," Natinal Health Agency, 2018. [Online]. E. K. Mario Arturo Ruiz Estrada, "The Evolution of Social Security Research,
5. " SSRC Working Paper Series, July 2017. S. C. Seng, "Social Security: Challenges and Issues,
6. " SSRC Working Paper Series No. 2014-1, January 2014E. K. Mario Arturo Ruiz Estrada, "The Evaluation of ASEAN-Members Pension Scheme Performance," Global Social Welfare, 2019. S. S. ,. A. A. Denis Garand, "Performance evaluationframework for government-sponsored health insuranceprogrammes," Background Paper,
7. Microinsurance network, 2015. S. Iyer, "Social Security Pension Scheme: An Analytical Approach, " https://doi.org/10.1017/S174849950000049X, 2008. L. Cao, "Social Security and Social Welfare Data Mining: An Overview," IEEE TRANSACTIONS ON SYSTEMS, MAN, and cybernetics—part c:
8. 19. applications and reviews, vol. 42, no. 6, 2012. H. Z. L. C. H. B. Y. O. C. Z. Yanchang Zhao, "DataMining Applications in Social Security,"in Domain Driven Data Mining, Springer, 2010, pp. 81-96.N. D. M. B. R. Maione, "Research on social data by means of Cluster Analysis,"
9. Applied Computing and Informatics, 2018. Y. F. W. Z. S. Z. Leibao Zhang, "
10. Subjective Well-Being Prediction Using Data Mining," Applied and Computational Mathematics, 2018. S. A. Wafa Qadadeha, "Customers Segmentation in
11. the Insurance Company (TIC) Dataset,"in INNS Conference on Big Data and Deep Learning ,
12. Edinburgh, UK, 2018. M. M. R. G. Seyed Mohammad Seyed Hosseini,
13. "Cluster analysis using data mining approach to develop CRM methodology," Expert Systems with
14. Applications, 2010. ⇑ . G. S. ,. M. H. ,. Y. L. ,. S. N. ,. D. C. C. Yang Xie a,
15. "Analyzing health insurance claims on different timescales to predict,
16. " Journal of Biomedical Informatics, 2016. H. P. J. L. W. C. J. Hyunjung Shin, "A scoring model to
17. detect abusive billing patterns in health insurance claims," Expert Systems with Applications, vol. 39, 2012. G. Vipula Rawte, "Fraud Detection in Health Insurance using Data mining Techniques," in International Conference on Communication, Information &
18. Computing Technology (ICCICT), Mumbai, 2015.]"https://backoffice. philipscapital.in",Philps capital, 2018[Online]

AUTHORS PROFILE

Sri P. Sunil Kumar completed his Master's degree in Computer Science and Applications in the year 1998.He is having 20 years of teaching experience and currently working as Assistant Professor in Sri Sathya Sai Institute of Higher Learning, Muddenahalli Campus and pursuing his PhD from Sri Sathya Sai Institute of Higher Learning, Puttaparthi. He has published 5 papers in the area of data mining. He is a member of CSI. He has guided more than 100 student projects at Undergraduate as well as at Postgraduate levels.

Dr Raghunatha Sarma.R has completed his PhD from Sri Sathya Sai Institute of Higher Learning in the year 2009.He has 16 research publications and completed 3 research projects namely Multi Resolution Based Image Mosaicing, Character Recognition for Indic Scripts. (Telugu, Tamil and Devanagari), and Ensembling of Image Segmentation Algorithms for Medical and Aerial Images, which are funded by DRDO and HP Labs.

Sri Satya Sai Mudigondais a Senior Tech Actuarial Consultant providing services in USA and India. He is leading Tech Actuarial projects involving high performance computing, block chain, machine learning, deep learning and other data science applications to insurance. He published twelve tech actuarial research papers in international journals. His significant contributions include developing framework for fraud detection, building dynamic ERM dashboards and using actuarial techniques as part of data science projects. He has guided number of student projects at post graduate level.