International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(1s) 2026 pp. 725-734
https://doi.org/10.70917/ijcisim-2026-1967
Article

Intelligent Recommendation Framework for Optimizing Risk in Medical Health Insurance Using Clustering and Association Rule Mining

Jyoti Navin Nandimath[1], Ganesh R. Pathak[2]

[1] MIT Art, Design and Technology University, Pune, Maharashtra, India.
Email: jyoti.nandimath@gmail.com
[2] MIT Art, Design and Technology University, Pune, Maharashtra, India.
Email: ganesh.pathak@mituniversity.edu.in

Abstract: The medical insurance industry is confronted with rising claim volatility and increasingly diverse customer profiles. Traditional actuarial techniques often struggle to capture subtle interactions among demographic, behavioral and clinical features, leading to coarse premium segmentation and sub-optimal risk sharing. This paper proposes an intelligent recommendation framework to optimize risk in medical health insurance. The approach uses a real claims dataset sourced from Kaggle, containing demographic, occupational and hereditary attributes. After cleaning duplicates, replacing missing values and applying Min–Max scaling and label encoding, exploratory analysis visualizes disease count distributions, hereditary conditions and job-title correlations. K-means clustering is employed to discover natural risk segments; optimum cluster numbers are determined using the silhouette and elbow methods. Cluster assignments are then used as input for association rule mining with Apriori and FP-Growth algorithms. Support, confidence and lift measures guide rule selection. The system ultimately produces rule-based recommendations for new applicants, predicting their likely claim group and suggesting premium strategies. Experiments demonstrate that seven clusters achieve a high silhouette score while seven clusters minimize within-cluster variation. FP-Growth is found to be faster and generates more concise rules than Apriori. The proposed framework aids insurance providers in understanding risk patterns, designing personalized products and reducing adverse selection. Future work will investigate deep learning and federated techniques to enhance scalability and privacy.

Keywords: Medical Health Insurance, Intelligent Recommendation System, Clustering Techniques, Association Rule Mining, Apriori and FP-Growth Algorithms, Risk.

1. Introduction

Personalized medical insurance has become a key differentiator in the ever-competitive insurance market. As populations are getting older and lifestyles have become more varied, insurers are forced to find a balance between treating their customers fairly and being profitable. Conventional premium determination mechanisms are based upon actuarial risk pools defined by very general characteristics such as age strata and disease history. These pools often ignore fine grained interrelations between variables such as the city of residence, hereditary diseases as well as occupational stressors and thus have healthy individuals subsidizing high risk customers, whilst individual high risk sub groups are under-priced. Advancements in machine learning (ML) and data mining provide opportunities to scrutinize large datasets of claims, and pick out complex patterns. By leveraging clustering together with the association rule mining technique, things that affect the value of insurers can be drawn. By leveraging clustering along with the association rule mining technique, homogeneous risk groups can be drawn and coherent recommendations can be produced[1], [2].

Modern recommendation systems in the healthcare field involve a broad range of fields. Some of the studies focus on optimizing return on investment (ROI) for insurers using artificial intelligence (AI) models, while others create personalized medical recommendation engines to aid in determining treatment or bring about healthy ageing.

Deep learning and graph clustering strategies have been used to implement temporally sensitive food recommendations, and ensemble learning has been successfully used in diabetes management. Numerous contributions highlight the need for domain expertise and specialized competency frameworks to enable engaging healthcare professionals to work with AI[3]. However, there remains a gap between general recommendation systems and insurance-specific risk optimization. Insurance datasets feature categorical fields such as job title, claim history and hereditary conditions. Simply training a predictive model on these labels may yield high accuracy but little interpretability. Customers and regulators demand transparent decision-making, particularly when premiums and coverage decisions are at stake. Rule-based recommendations derived from data mining address this need by providing explicit if-then clauses[4], [5].

This research builds upon earlier works in clustering and association rules. The paper considers a health insurance dataset comprising attributes such as age, gender, job title, city, hereditary diseases and claim amount. Duplicates are removed and missing values replaced with column means. Continuous features are scaled to the [0, 1] interval using the Min–Max transform and categorical fields are label-encoded. Exploratory visualizations reveal gender disparities in disease prevalence, the influence of hereditary conditions and correlations among variables[6], [7]. Using unsupervised clustering, we avoid imposing prior risk categories. Instead, clusters reveal latent segments that differ in claim behaviours. The silhouette and elbow methods guide the choice of cluster number. Once clusters are established, association rule mining extracts relationships between input conditions and claim outcomes. For example, a rule might state that "female, hereditary disease and cluster 3 → high claim group" with a confidence of 75 %. Such rules are intuitive and actionable for underwriters.[8]

The existing works shows promising advances in personalized recommendation and clustering techniques. Yet, few studies integrate unsupervised clustering with rule mining specifically for insurance risk. Our contribution lies in coupling K-means clustering with Apriori and FP-Growth algorithms on cluster-labelled data to produce transparent recommendations. We design metrics to select optimal cluster counts and evaluate rule quality using support and confidence thresholds. The objectives of the proposed research are:

i. To confirm domain-specific performance metrics by conducting experiments on a real insurance dataset and assessing clustering quality and rule effectiveness.

ii. To propose an intelligent recommendation algorithm that combines clustering and association rule mining to optimize medical insurance risk for service providers.

The rest of the paper is organized as follows: Section 2 deals with the review of existing work, Section 3 details the methodology, including data preprocessing, exploratory analysis, clustering and rule mining. Section 4 discusses results and interprets figures. Section 5 concludes and outlines future research directions.

2. Literature Review

Recent research in healthcare recommendation systems and insurance analytics spans AIdriven ROI optimization, personalized medicine, smart healthcare platforms and timeaware food recommenders. For example, a 2024 study on smart healthcare systems combines data preprocessing, clustering and performance evaluation to derive insights from medical records. Another work delivers a personalized medical recommendation system using support vector classifiers and random forests with high accuracy. A 2022 paper develops a timeaware food recommender using deep learning and graph clustering to address cold start and temporal effects. Collectively, these works demonstrate the value of clustering and AI in personalized recommendations.

Author(s) & Year	Methodology	Domain/ Application	Key Finding	Algorithm/Model	Data Source
Sellamuthu et al. [9] (2023)	Hybrid AI-based recommender combining supervised learning & optimization	Insurance ROI maximization	AI models improve investment decisions & ROI	Neural networks & optimization	Simulated insurance data
Hassan et al.[10](2025)	Personalized recommendation using machine learning	Medical advice	SVC & RF achieve high disease-prediction accuracy	SVC, Random Forest	Symptom–disease dataset
Martinez-Rodrig	Data analysis &	Personalized	Integrating	Statistical	Vascular

o et al.[11] (2024)	recommendation for vascular ageing	medicine	recommendations with health data supports healthy ageing	methods & recommender	health records
Russell et al.[12] (2023)	Competency framework development	Healthcare workforce	Identified essential AI skills for healthcare professionals	Surveys & qualitative analysis	Workforce surveys
Yang et al.[13] (2024)	Clustering & data analysis (SHCM)	Smart healthcare	Combining SOM mapping & K-means improves clustering performance	SOM + K-means	Outpatient clinic database
Kumar et al.[14] (2024)	Interactive recommender using big data analytics	Health recommendations	Big-data pipelines enable interactive healthcare advice	Hadoop & ML algorithms	Health logs
Vullam et al.[15] (2023)	Multi-agent personalized recommender	E-commerce	Agents tailor product recommendations	Multi-agent system & collaborative filtering	Online shopping data
Rostami et al.[16] (2022)	Time-aware food recommender	Nutrition & diet	Deep learning & graph clustering mitigate cold start issues	LSTM & graph clustering	User diet logs
Hassan et al.[17] (2022)	Review of genomics & big data	Personalized medicine	Big data analytics revolutionizes personalized medicine	Literature synthesis	Genomic & big data sources
Ihnaini et al.[18] (2021)	Deep ensemble learning	Diabetes management	Ensemble models enhance multidisciplinary recommendations	Deep neural networks	Multidisciplinary health records
Karthik et al.[19] (2021)	Fuzzy recommender using sentiment & ontology	E-commerce	Fuzzy logic & sentiment analysis predict customer interests	Fuzzy logic & ontology	Review texts & sentiments

Existing studies address a variety of recommendation problems—ranging from ROI optimization and personalized medical advice to smart health systems and timeaware food recommendations—yet few focus specifically on optimizing healthinsurance risk. Many systems either apply supervised models with limited interpretability, or examine clustering without linking it to actionable recommendations. Our research bridges this gap by coupling Kmeans clustering with rule mining to produce transparent insurancerisk recommendations. It emphasizes interpretability, enabling insurers to understand why certain customers fall into highrisk segments, and offers a novel, domainspecific contribution to personalized insurance analytics.

3. Methodology

The proposed framework follows a sequential pipeline: data preprocessing, exploratory data analysis (EDA), clustering to discover risk segments and association rule mining to build recommendations. Each stage involves mathematical definitions and design choices.

3.1. *Data Preprocessing*

The dataset contains numeric attributes (e.g., age, claim amount) and categorical fields ("*sex, city, hereditary disease, claim group and job title*")[20]. Duplicate rows are removed to avoid biases. Let denote the original dataset with samples and features. For each feature , missing values are imputed using the column mean shown in eq.1:

$$\tilde{x}_{ij} = \begin{cases} x_{ij} & if\ x_{ij} \neq missing, \\ \frac{1}{n_j \sum_{i:x_{ij} \neq missing} x_{ij}} & otherwise \end{cases} \qquad (1)$$

where is the number of nonmissing observations in feature j. After imputation, numeric attributes are normalized using Min–Max scaling to bring them into the [0,1] interval shown in eq.2:

$$x_{ij}^{norm} = \frac{x_{ij} - \min_i(x_{ij})}{max_i(x_{ij}) - min_i(x_{ij})} \qquad (2)$$

Categorical variables are transformed into integer codes via label encoding . This yields a clean feature matrix suitable for clustering algorithms.

	age	sex	weight	bmi	hereditary_diseases	no_of_dependents	smoker	\
0	60.0	male	64	24.3	NoDisease	1	0	
1	49.0	female	75	22.6	NoDisease	1	0	
2	32.0	female	64	17.8	Epilepsy	2	1	
3	61.0	female	53	36.4	NoDisease	1	1	
4	19.0	female	50	20.6	NoDisease	0	0	

	city	bloodpressure	diabetes	regular_ex	job_title	claim
0	NewYork	72	0	0	Actor	13112.6
1	Boston	78	1	1	Engineer	9567.0
2	Phildelphia	88	1	1	Academician	32734.2
3	Pittsburg	72	1	0	Chef	48517.6
4	Buffalo	82	1	0	HomeMakers	1731.7

Figure 1 Sample dataset

	age	sex	weight	bmi	hereditary_diseases	no_of_dependents	smoker	city	bloodpressure	diabetes	regular_ex	job_title	claim
0	60.0	male	64	24.3	NoDisease	1	0	NewYork	72	0	0	Actor	13112.6
1	49.0	female	75	22.6	NoDisease	1	0	Boston	78	1	1	Engineer	9567.0
2	32.0	female	64	17.8	Epilepsy	2	1	Phildelphia	88	1	1	Academician	32734.2
3	61.0	female	53	36.4	NoDisease	1	1	Pittsburg	72	1	0	Chef	48517.6
4	19.0	female	50	20.6	NoDisease	0	0	Buffalo	82	1	0	HomeMakers	1731.7

Figure 2 Dataset after set of preprocessing

3.2. *Exploratory Data Analysis*

EDA is performed to understand the distribution of diseases, hereditary conditions and job titles. Figure-3 illustrate that male employees have a slightly higher disease count than females, while hereditary diseases are concentrated in a few job categories. A grouped bar chart compares disease incidence across job categories. Physically demanding jobs have higher disease counts than sedentary ones, implying that occupational risk influences claim likelihood

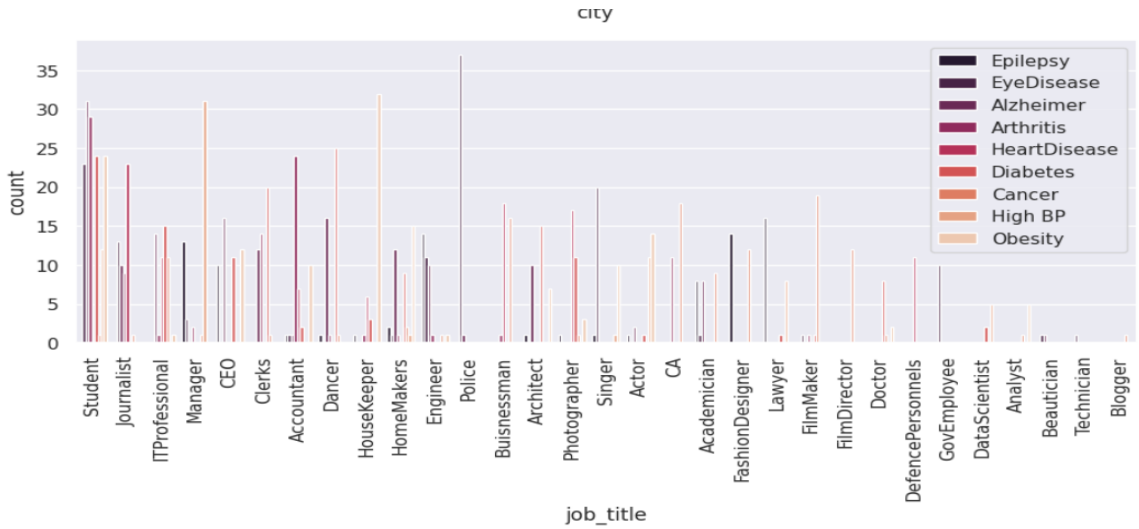

Figure 3 Job Title vs Disease Count

The correlation matrix as shown in figure-4 highlights that claim amounts correlate moderately with age and job title, motivating the use of multivariate analysis.

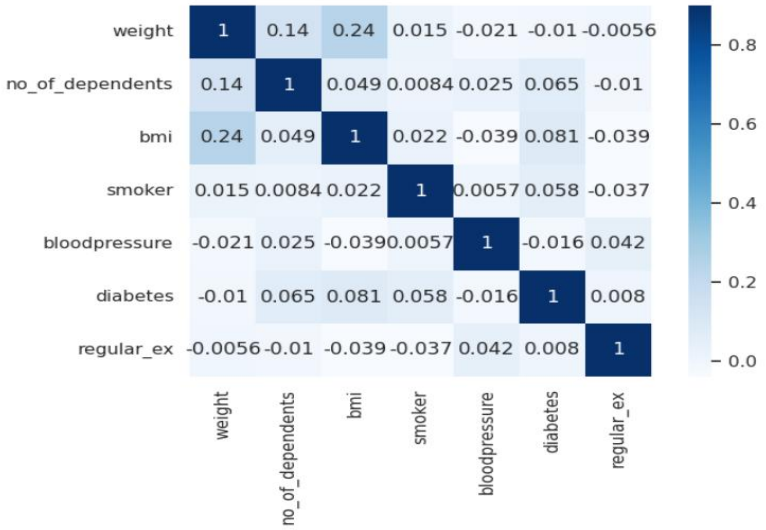

Figure 4 Grouping of Claim column data

3.3. Clustering Techniques

Kmeans clustering partitions samples into clusters by minimizing the "*withincluster sum of squares*" (WCSS). The objective function is represented as in eq.3:

$$WCSS(K) = \sum_{k=1}^{K} \sum_{i \in C_k} ||X_i - \mu_k||^2 \qquad (3)$$

where is a sample in cluster with centroid . The optimal number of clusters is unknown a priori. Two heuristic methods help determine K:

- The silhouette coefficient, defined for sample as in eq.4

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \qquad (4)$$

- where is the average distance from to other points in its cluster and is the smallest average distance to points in another cluster. Values near 1 indicate wellclustered samples. The mean silhouette across all samples guides the choice of . Figure-5, 6 plot silhouette scores versus , showing a maximum at K=7 as shown in figure-7.

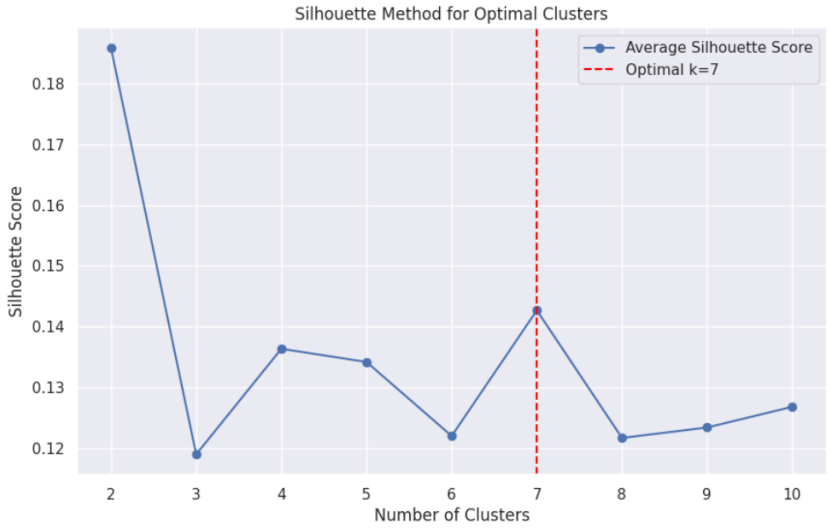

Figure 5 Optimal Cluster Finding using Silhouette Method

```
For k=2: Min Silhouette Score = -0.0583, Max Silhouette Score = 0.3128
For k=3: Min Silhouette Score = -0.0756, Max Silhouette Score = 0.3013
For k=4: Min Silhouette Score = -0.0477, Max Silhouette Score = 0.4214
For k=5: Min Silhouette Score = -0.0596, Max Silhouette Score = 0.4226
For k=6: Min Silhouette Score = -0.0718, Max Silhouette Score = 0.4238
For k=7: Min Silhouette Score = -0.0488, Max Silhouette Score = 0.4248
For k=8: Min Silhouette Score = -0.0710, Max Silhouette Score = 0.4226
For k=9: Min Silhouette Score = -0.0982, Max Silhouette Score = 0.4325
For k=10: Min Silhouette Score = -0.0911, Max Silhouette Score = 0.4316
```

Figure 6 Cluster and their respective Silhouette score

	age	sex	weight	bmi	...	job_title	claim_group	Cluster
14990	0.543478	1	0.622951	0.539084	...	28	12	5
14991	0.500000	0	0.655738	0.156334	...	22	6	4
14992	0.695652	0	0.426230	0.274933	...	20	6	6
14993	0.826087	0	0.344262	0.261456	...	12	6	4
14994	0.434783	0	0.606557	0.304582	...	10	5	3
14995	0.456522	1	0.245902	0.331536	...	20	8	5

Figure 7 Clustering with optimal cluster value (k=7)

- The elbow method plots WCSS(K) against KKK and looks for a “knee” where additional clusters yield diminishing improvement. Figure-8 indicates an elbow near K=4. Both methods are evaluated to compare segmentation quality.

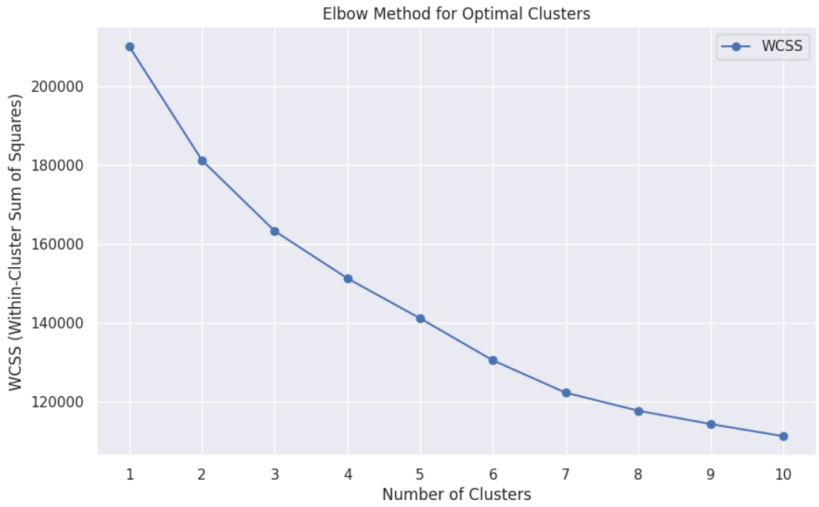

Figure 8 Optimal Cluster Finding using Elbow Method

After selecting K, Kmeans assigns each sample to the nearest centroid. Cluster membership labels are appended to the dataset and saved as cluster.csv. Visualization using principal component analysis (PCA) and tdistributed stochastic neighbor embedding (tSNE) projects highdimensional data into 2D/3D for inspection as shown in figure-9. These projections reveal compact clusters and overlapping boundaries, confirming segmentation validity.

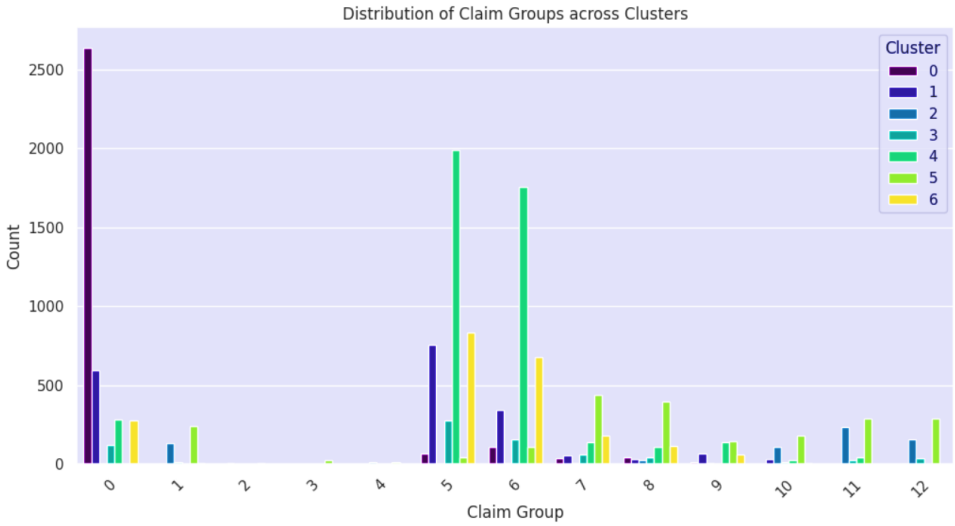

Figure 9 Claim Groups across various cluster

3.4. Association Rule Mining

Once cluster labels are created, the task becomes discovering rules that relate demographic attributes and cluster assignments to claim groups. Consider an itemset III comprising attribute–value pairs such as . The support of is represented in eq.5:

$$supp(I) = \frac{total\ number\ of\ transactions}{number\ of\ transactions\ containing\ I} \qquad (5)$$

An association rule takes the form , where are disjoint itemsets. The confidence of the rule is represented in eq.6,7:

$$conf\,(I_X \Rightarrow I_Y) = \frac{supp(I_X \cup I_Y)}{supp(I_X)} \qquad (6)$$

And the lift is

$$lift(I_X \Rightarrow I_Y) = \frac{conf(I_X \Rightarrow I_Y)}{supp(I_Y)} \quad (7)$$

Here implemented the Apriori algorithm, which generates candidate itemsets iteratively using the property that any subset of a frequent itemset must also be frequent. For each cluster, Apriori extracts rules where the antecedent includes demographic attributes and the consequent is a claim group label. FPGrowth, an alternative algorithm, builds a compressed prefix tree (FPtree) and generates frequent patterns without candidate generation, resulting in faster execution on sparse data [8]. Minimum support and confidence thresholds are tuned experimentally.

3.5. Prediction and Recommendation

For a new applicant with demographic information x, the system determines the nearest cluster using the trained Kmeans centroids. Let c denote the predicted cluster. The recommendation engine searches the set of rules extracted for cluster c. Given antecedent conditions matching the applicant, the consequent claim group is returned along with its confidence. The insurer can interpret this as a probability of a high claim and adjust premiums accordingly. If multiple rules apply, the one with the highest lift or confidence is chosen.

3.6. Evaluation Metrics

Clustering performance is evaluated using the average silhouette score and the ratio of intracluster to intercluster distances. Rule mining quality is assessed via the number of rules, their average support and confidence. Comparative analysis between Apriori and FPGrowth reveals tradeoffs in speed and rule comprehensibility. The framework's effectiveness is measured by its ability to correctly predict claim groups on a heldout test set and provide actionable recommendations.

4. Results And Outputs

4.1. Clustering Technique

The 3D PCA plot is used to display the spatial distribution of the data points in the seven different clusters after the data has been dimensionally reduced. Each cluster represents a distinct configuration in the feature space, implying that individuals placed in the same cluster have similar insurance related characteristics. The delineation that is apparent along PCA components 1 and 2 verifies that the clustering algorithm is separating customer segments based upon underlying risk attributes as shown in figure-10, 11. The dense clustering of points supports an existence of a well-defined structure in the dataset making it suitable for segmentation. As a result, this visual representation justifies the substantive relevance of the clustering stage before the association rule mining for recommendation generation can be deployed.

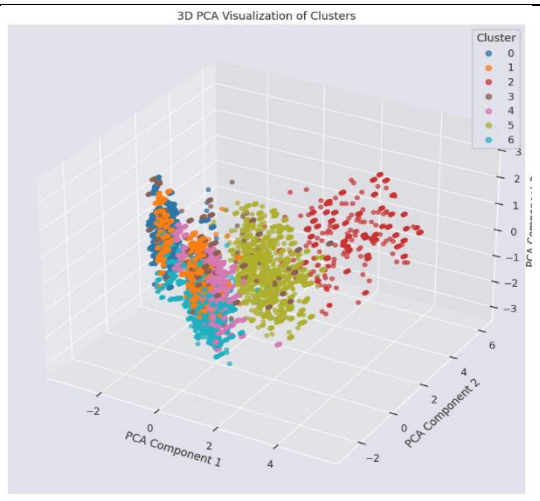

Figure 10 Cluster visualization Using PCA

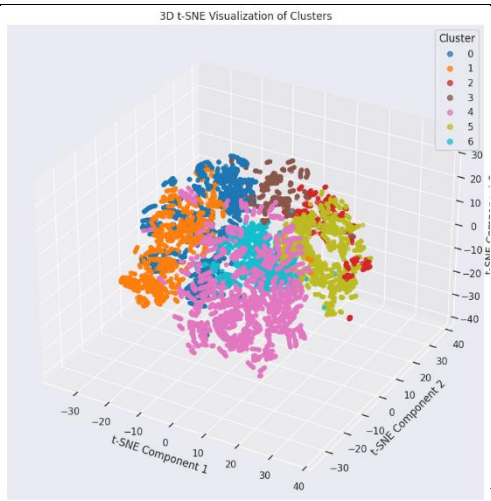

Figure 11 Cluster Visualization Using t-SNE

The t-SNE visualization illustrates the separation between clusters even more than linear dimensionality reduction methods are able to do. In comparison to PCA, t-SNE shows tighter borders and an increased contrast between clusters, thus favoring the formation of compact groupings. This phenomenon shows that claim behavior and risk attributes have non-linear relationships that are well captured using t-SNE. The greater isolation from clusters indicates that there is a greater consistency within each predicted cohort. These results support the clustering as a valid base for the derivation of insurance risk patterns and the provision of personalized policy advice.

4.2. Association Rule Mining

The comparative assessment chart is based on the assessment of Apriori and FP- Growing algorithms with respect to execution time, number of frequent itemsets and number of generated rules. It can be seen from the execution time measure that FP-Growth achieves much better performance than Apriori, especially when the data set size becomes large, thanks to its tree-based compression mechanism. While both algorithms discover a similar number of frequent itemsets and rules, the better processing efficiency of the FP-Growth algorithm emphasizes its suitability for decision support environments that require real-time decision making as shown in figure-12. The near equivalence in the number of rules confirms that the accelerated performance does not affect the accuracy

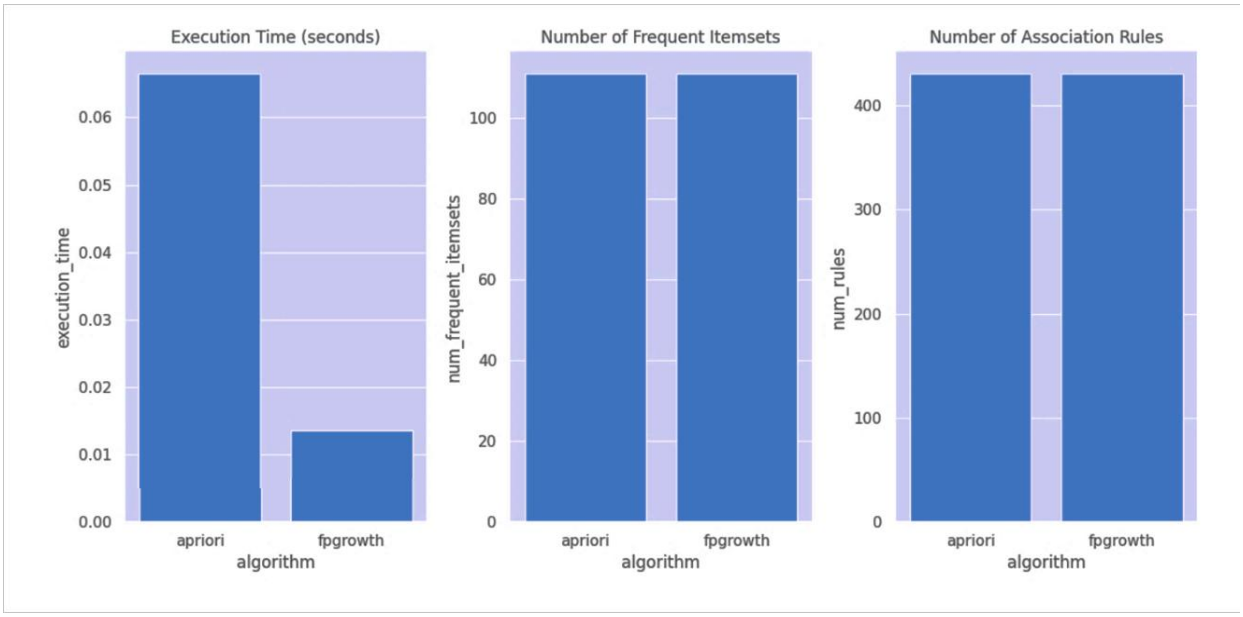

Figure 12 Comparative Analysis of Association Rule Algorithm

5. Conclusion And Future Scope

The study introduces an intelligent recommendation framework that combines Kmeans clustering with association rule mining to optimize medical insurance risk. By cleaning and normalizing a real healthinsurance dataset, applying exploratory analysis, determining optimal clusters via silhouette and elbow methods and extracting rules using Apriori and FPGrowth, the approach uncovers latent risk segments and generates transparent recommendations. Experiments show that seven clusters provide high silhouette scores; FPGrowth produces faster, more concise rules than Apriori. The system developed herein provides insurers with the ability to tailor premiums, coverages, according to individual risk profile which will reduce instances of adverse selection and provide a fairer system.

This study helps bridge the divide between the more generic recommendation systems and insurance specific risk optimization. It shows that unsupervised clustering with association rule mining is an effective combination to help us produce understandable knowledge based on heterogeneous medical data. In addition, the explanations generated by the algorithm, which are based on rules, meet regulatory requirements for transparency and help to build confidence at the decision maker level.

Future research may extend the framework by using deep learning methods to learn richer representations, incorporating federated learning methods to preserve privacy, supervised classifiers to improve predictive accuracy

and external data-sources like electronic health record or lifestyle tracker. Such improvements would support a more whole-of-picture and proactive approach to managing risks.

References

1. G. J. Oyewole and G. A. Thopil, “Data clustering: application and trends,” Artif. Intell. Rev., vol. 56, no. 7, pp. 6439–6475, 2023, doi: 10.1007/s10462-022-10325-y.
2. T. N. T. Tran, A. Felfernig, C. Trattner, and A. Holzinger, “Recommender systems in the healthcare domain: state-of-the-art and research issues,” J. Intell. Inf. Syst., vol. 57, no. 1, pp. 171–201, 2021, doi: 10.1007/s10844-020-00633-6.
3. W. Li et al., “A Comprehensive Survey on Machine Learning-Based Big Data Analytics for IoT-Enabled Smart Healthcare System,” Mob. Networks Appl., vol. 26, no. 1, pp. 234–252, 2021, doi: 10.1007/s11036-020-01700-6.
4. S. A. Alowais et al., “Revolutionizing healthcare: the role of artificial intelligence in clinical practice,” BMC Med. Educ., vol. 23, no. 1, p. 689, 2023, doi: 10.1186/s12909-023-04698-z.
5. D. Roy and M. Dutta, “A systematic review and research perspective on recommender systems,” J. Big Data, vol. 9, no. 1, p. 59, 2022, doi: 10.1186/s40537-022-00592-5.
6. M. Badawy, N. Ramadan, and H. A. Hefny, “Healthcare predictive analytics using machine learning and deep learning techniques: a survey,” J. Electr. Syst. Inf. Technol., vol. 10, no. 1, p. 40, 2023, doi: 10.1186/s43067-023-00108-y.
7. H. Ko, S. Lee, Y. Park, and A. Choi, “A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields,” Electronics, vol. 11, no. 1. p. 141, 2022, doi: 10.3390/electronics11010141.
8. R. De Croon, L. Van Houdt, N. N. Htun, G. Štiglic, V. Vanden Abeele, and K. Verbert, “Health Recommender Systems: Systematic Review,” J Med Internet Res, vol. 23, no. 6, p. e18035, 2021, doi: 10.2196/18035.
9. S. Sellamuthu et al., “AI-based recommendation model for effective decision to maximise ROI,” Soft Comput., 2023, doi: 10.1007/s00500-023-08731-7.
10. B. M. Hassan and S. M. Elagamy, “Personalized medical recommendation system with machine learning,” Neural Comput. Appl., vol. 37, no. 9, pp. 6431–6447, 2025, doi: 10.1007/s00521-024-10916-6.
11. A. Martinez-Rodrigo, J. C. Castillo, A. Saz-Lara, I. Otero-Luis, and I. Cavero-Redondo, “Development of a recommendation system and data analysis in personalized medicine: an approach towards healthy vascular ageing,” Heal. Inf. Sci. Syst., vol. 12, no. 1, p. 34, 2024, doi: 10.1007/s13755-024-00292-9.
12. R. G. Russell et al., “Competencies for the Use of Artificial Intelligence–Based Tools by Health Care Professionals,” Acad. Med., vol. 98, no. 3, 2023, [Online]. Available: https://journals.lww.com/academicmedicine/fulltext/2023/03000/competencies_for_the_use_of_artificial.19.aspx.
13. W.-C. Yang, J.-P. Lai, Y.-H. Liu, Y.-L. Lin, H.-P. Hou, and P.-F. Pai, “Using Medical Data and Clustering Techniques for a Smart Healthcare System,” Electronics, vol. 13, no. 1. p. 140, 2024, doi: 10.3390/electronics13010140.
14. M. S. Kumar et al., “An Interactive Healthcare Recommendation System Using Big Data Analytics,” in 2024 3rd International Conference for Advancement in Technology (ICONAT), 2024, pp. 1–6, doi: 10.1109/ICONAT61936.2024.10774650.
15. N. Vullam, S. S. Vellela, V. R. B, M. V Rao, K. B. SK, and D. R., “Multi-Agent Personalized Recommendation System in E-Commerce based on User,” in 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2023, pp. 1194–1199, doi: 10.1109/ICAAIC56838.2023.10140756.
16. M. Rostami, M. Oussalah, and V. Farrahi, “A Novel Time-Aware Food Recommender-System Based on Deep Learning and Graph Clustering,” IEEE Access, vol. 10, pp. 52508–52524, 2022, doi: 10.1109/ACCESS.2022.3175317.
17. M. Hassan et al., “Innovations in Genomics and Big Data Analytics for Personalized Medicine and Health Care: A Review,” International Journal of Molecular Sciences, vol. 23, no. 9. p. 4645, 2022, doi: 10.3390/ijms23094645.
18. B. Ihnaini et al., “A Smart Healthcare Recommendation System for Multidisciplinary Diabetes Patients with Data Fusion Based on Deep Ensemble Learning,” Comput. Intell. Neurosci., vol. 2021, no. 1, p. 4243700, Jan. 2021, doi: https://doi.org/10.1155/2021/4243700.
19. R. V Karthik and S. Ganapathy, “A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce,” Appl. Soft Comput., vol. 108, p. 107396, 2021, doi: https://doi.org/10.1016/j.asoc.2021.107396.
20. S. Gupta, “Health insurance data set.” www.kaggle.com, 2021, [Online]. Available: https://www.kaggle.com/datasets/sureshgupta/health-insurance-data-set/data.