Technological Forecasting & Social Change 201 (2024) 123217

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Developing strategies to retain organizational insurers using a clustering technique: Evidence from the insurance industry

Kambiz Shahroodi [a], Soroush Avakh Darestani [b,*], Samaneh Soltani [a], Adeleh Eisazadeh Saravani [a]

[a] Department of management, Islamic Azad University, Rasht Branch, Rasht, Iran
[b] Guidhall School of Business and Law, London Metropolitan University, London, United Kingdom

ARTICLE INFO

Keywords:
Influential factors on churn
Insurance industry
Forecasting
Clustering
K-means technique

ABSTRACT

Formulating strategies to maintain policyholders is one of the main challenges for most insurance companies in Iran. The purpose of this article is to help marketing strategists of insurance companies predict insurees' churn and develop insurees retention strategies. Since the cost of maintaining an insurance policyholder is approximately one-eighth of the cost of attracting new ones, predicting their churn can help insurance companies adopt proper strategies in advance, which will definitely lead to saving marketing costs and maintaining the insurer's portfolio. Accordingly, the main question of this research is how to classify organizational insurees with the help of the clustering technique. This research is conducted in both qualitative and quantitative phases. In the qualitative phase, by conducting a semi-structured interview (interview protocol) with 15 experts in the insurance industry, the influential factors on policyholders' churn are identified. Then, based on the factors identified in the research literature and comparing them with the interview results, eight main factors are finalized. In the quantitative phase, in order to cluster the organizational insurees, 120 samples from the Iran Insurance Company are selected, and k-means is applied for clustering. Organizational insurees are divided into two groups according to the desired indicators. Using the results of clustering, insurees are divided into four groups, and effective marketing strategies are developed for each group. According to the results, the variable "health care insurance price" has the most effective role in separating the clusters at an error level of <0.01, and on the contrary, the variable "liability insurance amount" has the least important role at an error level of <0.978.

1. Introduction

Customers are one of the most important assets of businesses in many dynamic and competitive companies (Amin et al., 2018). As existing customers are an important source of profit in business, the ability to identify customers who show signs of leaving the company can generate more profit for companies (Chen, 2016). In the era of increasingly saturated markets and intensified competition among companies, customer churn has become a fundamental problem. Searching and identifying customers who show a high tendency to leave the company or, in other words, predicting customer churn, is of significant importance in a customer retention strategy that aims to reduce customer churn (De Caigny et al., 2018). Customer churn is a marketing term which occurs when customers switch to other suppliers or buy less. Managing customer churn is one of the major challenges that organizations face, especially those offering subscription-based services (Prabadevi et al., 2023).

When customer expectations are not met, the opposite can occur, i.e., customer churn. Customer churn is the loss of an existing customer to a competitor. In this research a competitor is a different brand, which can result in a churning customer although the customer stays at the same company. To manage customer churn, the churning customers should be recognized and then these customers should be induced to stay (Huigevoort, 2015).

Customer churn, or the propensity at which clients cease their business relationship with a company, poses a big challenge in various industries not exempting insurance. Losing customers leads to financial loss because of reduced sales and it further leads to an increasing need for attracting new customers. As confirmed by related studies, reducing attrition rate and focusing on existing policyholders are deemed as the more efficient and cost-effective marketing approach which maximizes shareholder's value. Long term customers would be more beneficial and,

* Correspondence author at: Guidhall School of Business and Law, London Metropolitan University, London, United Kingdom.
E-mail address: s.avakhdarestani@londonmet.ac.uk (S. Avakh Darestani).

https://doi.org/10.1016/j.techfore.2024.123217
Received 19 March 2023; Received in revised form 4 January 2024; Accepted 5 January 2024

if satisfied, may provide new referrals. Hence, customer churn retention analyses are designed to predict which customers are about to churn and facilitate an accurate segmentation of the market which allows organizations to target the customers who are most likely to churn with a retention campaign. With effective churn management, losses due to churn are minimized through prediction and profits are maximized by retaining valuable customers (Bravante and Robielos, 2022). In today's competitive environment, customers tend to easily compare and switch between service providers. Customers who decide to leave are generally named as churned customers (Liu et al., 2024). Günther et al. (2014) defined that the customer has churned "When a customer cancels all his/her policies, either to switch insurance provider or because the need of insurance is no longer present". However, there is a differentiation in the types of customers' churns, and not all are suitable for analysis. A churner can be classified as voluntary when the customer deliberately decides to switch to another service provider or involuntary (usually not included for modeling) if circumstances like death or bankruptcy occur.

Due to limited differentiation among service providers in competitive markets in insurance industry, customers can easily compare other insurance providers' offers in terms of premiums and coverage just using the internet. As a result, this transparency leads many customers to voluntarily switch to other insurance providers and cancel the insurance product they have on the company. This transparency has a significant impact on the insurance market's competitiveness. Given this, especially now, to enhance competitiveness, companies should focus on reducing customer churn. Although Reichheld and Sasser, (1990) are the grounders of the "zero defections" theory, they defend that companies should not try to eliminate all defections but be prepared to spot customers who leave and act accordingly to their findings (Castro, 2022).

According to the statistics of the Central Insurance of Iran, the rate of churn in 2022 had a 15 % decrease, compared to 2021, in Iran's insurance industry. This led to a drop of 3000 billion Tomans in insurance premiums. This means that effective marketing and customer maintenance strategies could be opted for if the churn had been predicted in advance (Central Insurance of Iran, 2022). In recent years, the insurance industry in Iran has undergone many fluctuations in terms of profitability, portfolio composition, loss rate, penetration rate, retention and satisfaction of policyholders, and market share as a result of the emergence of numerous companies in the competitive market. Retention of policyholders has become the main problem for most insurance companies. Since in Iran's insurance industry, like any other industry, the cost of searching for new insurees is much more than the cost of maintaining current ones, it is necessary to analyze the reasons for insurees' churn in this industry and design models to predict the number of insurees who will leave the company's portfolio in the coming years. There should be models to enable predicting churn in the coming years. The identification of these factors lets insurance companies prepare themselves in advance to deal with this problem. Preliminary studies in Iran's insurance industry show that policyholders churn for several reasons; the main reasons are failure to receive claims on time, high rates of insurance premiums, improper behavior of employees of insurance companies, unfair activities of competing insurance companies, evasion of insurance companies from paying claims, delay on the part of insurance companies to provide services after the sale of insurance policies, etc. Consequently, the main problem of Iran's Insurance Industry is the high rate of policyholders' churn. Therefore, the main research questions (RQ) are formulated as follows:

- RQ 1: What are *the most important factors* influencing the churn of insurees in Iran's insurance industry?
- RQ 2: Is it possible to design a model, with the help of statistical and mathematical techniques, to *categorize* insurees in Iran's insurance industry?
- RQ 3: Is it possible to design a model, with the help of statistical and mathematical techniques, to *predict* their churn?

1.1. Contribution of research

Categorizing and clustering organizational insurees into two groups: churned or retained, will help insurance companies formulate and implement more effective marketing strategies for policyholders who are more likely to churn. Therefore, the clustering of insurees as a technique can lead to the design of more effective marketing strategies. The contribution of this work is to design a hybrid neural network and regression model associated with decision-making techniques to predict the churn of organizational insurance policyholders. Hence, organizational insurees are analyzed, and the indicators used in the model are specific to organizational insurees; namely: the nature of the insured and the age of the company. Moreover, the combined technique of neural networks and clustering is used to classify and predict the churn in institutional insurees.

According to the contribution of research, this study explores the utilization of a hybrid model for predicting the churn of organizational insurance policyholders. The remainder of this paper is organized as follows: Section 2 introduces related work on variables in insurance policies and CCP. Section 3 presents the research methodology. The results are discussed in Section 4, and finally, the conclusions are presented in Section 5.

2. Literature review

Customer churn poses a big challenge for financial institutions, including in the field of insurance. Hence, retention analyses are designed to predict time to churn and identify which among the customers are most likely to leave the company so that retention campaigns and strategies may be done to address the issue (Bravante and Robielos, 2022). Customer churn has been increasing in insurance, mainly due to technological improvements that allow customers to explore other insurance providers' offers. Given this, insurance providers need to compete among them, not only to get new customers but also to maintain their own (Castro, 2022). The first section of the literature review is assigned to reviewing research that identified and addressed the main variables of insurance policies.

2.1. Related work on variables of insurance policies

Leiria et al. (2022) identified the main factors that explain the cancellation of motor insurance policies by individual customers, considering the influence of intermediaries on their decisions. They found that aggressive tactics by insurance companies for customer acquisition may induce the cancellation of insurance policies. More valuable customers, policies with higher premiums and recent claims, as well as ancillary intermediaries and agents, are determinants of insurance cancellation. Conversely, the payment of policies by direct debit or without installments reduces the probability of cancellations (Leiria et al., 2022).

In this context, another study was conducted by Chen et al. (2022). They proposed a new model, i.e., combined the Cox model with variable penalties to model the clients' churn problems and determine the crucial factors that affect clients' decisions based on their personal information and behavior data, which are provided by a large insurance company in China. This model proved to be successful in identifying the client churning and making comparisons among the penalties. The variable penalties model reveals the most important factor about clients' churn problems, which can provide a reliable basis for the product development of insurance companies (Chen et al., 2022).

Matthijs Verschuren (2022) believes that insurers are increasingly adopting more demand-based strategies to incorporate the indirect effect of premium changes on their policyholders' willingness to stay. Matthijs Verschuren (2022) considered a causal inference approach in his work to account for customers' price sensitivity and to deduce optimal, multi-period profit-maximizing premium renewal offers

(Matthijs Verschuren, 2022). In this context, Soltani Lifshagerd et al. (2021) conducted research that aimed to find the main factors of churn in the Iranian insurance industry. From a contribution perspective, they proved that the eight identified factors (type of insurance, premium, final result of claims, duration of cooperation, payment method, number of installments, number of policies, and number of claims) can best explain the reasons for the churn of policyholders (Soltani Lifshagerd et al., 2021).

2.2. Related work on customer churn prediction

This section discusses the literature related to customer churn prediction (CCP). Churn prediction is one of the most crucial stages in customer relationship management (CRM).

The acquisition cost is the price an enterprise pays to gain new clients. On the other hand, retention costs are the costs of maintaining existing clients. It is actually difficult to predict which customers will churn and which customers will be maintained due to human limitations. The acquisition cost is approximately five times that of the retention cost (Wagh et al., 2024). Therefore, developing a precise and effective customer churn prediction model is a key factor in customer retention. In recent years, several studies have been carried out using static features, especially in sectors with large numbers of customers, such as telecommunications and insurance (Xing et al., 2022). Bellani (2019) stated that the generation of churn models, also known as retention or attrition models, is a growing problem in many industries. In the insurance industry, these customers cancel their contract or policy to benefit from better conditions (a lower premium) with another company. In her research entitled "Predictive Churn Models in Vehicle Insurance," The contribution of the work was focusing on vehicle policies, and the goal of her project was to develop a predictive model to reduce customer churn from a company. The predictive models employed were generalized linear models and artificial neural networks; parameter tuning was also conducted (Bellani, 2019).

To predict customer churn, Bolancé et al. (2016) focused on a real-life case in the motor insurance sector. The main question of their work is how to predict churn as the fundamental factor in the prevention of revenue loss. They proposed four different methods to predict lapsing from a portfolio of policies. They presented a comparative analysis between three different performance measures to assess the predictive power of each model. Their comparison analyzed the outcomes of a logistic regression, a conditional tree, a neural network, and a support vector machine. These are all considered basic approaches to data mining and knowledge discovery. The main contribution of their research was to show that, depending on the type of analysis and the research objective, the optimal prediction method may differ (Bolancé et al., 2016).

One way of predicting customer churn is by employing machine learning algorithms. In this context, another study was conducted by Prabadevi et al. (2023) using machine learning algorithms for customer churn prediction. The main question of their work was how to advise on the optimum machine-learning strategy for early client churn prediction. They proposed a system involved in the analysis of customer churning that included four different algorithms: stochastic gradient booster, random forest model, K-Nearest Neighbors, and logistic regression model (Prabadevi et al., 2023). Moreover, Wagh et al. (2024) employed machine learning methods to predict customer churn in the telecom sector.

Findings from existing studies have indicated that the churn rate is normally measured for a specific time window. Primarily, organizations motives should be toward customers' satisfaction and retaining their existing customers. In fact, retaining existing customers is equally important as absorbing new ones. It is noteworthy to say that customer churn prediction is the most important factor in adopting an industry's product or service Prabadevi et al. (2023). Given the importance of increasing the insurance penetration rate in Iran, which is currently 2.3 (according to Iran's central insurance statistics in 2021), and comparing it with the average insurance penetration rate in the world, which is about 9, it is significant to look for ways to improve this rate. One of the alternatives is to identify the factors affecting the churn of insurance policyholders and design a model to predict it. Further research should be conducted by including an adaptive learning approach for CCP analysis in the insurance industry and services. (Amin et al., 2023). Recently, another study conducted by Liu et al. (2024) applied extreme gradient-boosting trees to predict profit-driven customer churn. They used real datasets from service providers in multiple markets and found that marketing experts can design targeted marketing plans to maintain customer groups with a higher likelihood of churning. In this context, Usman-Hamza et al. (2024) conducted an empirical analysis of the effectiveness of tree-based machine-learning classifiers with different computational characteristics in the presence of a class imbalance problem for CCP.

In summary, according to the reviewed literature and studies compared in Table 1, a few studies have been conducted regarding CCP in the insurance sector. In addition, there are not studies that benefit and associate decision-making and regression models with neural networks to develop strategies to retain organizational insurees in the insurance industry. Hence, this study attempts to develop a hybrid framework for developing strategies to maintain insurees especially in the insurance sector. Table 1 is a summary of relevant studies in terms of the CCP, methods and data, industry, and contribution to address the gap in research.

3. Research methodology

The research framework includes all the required steps for this work. The framework for this study is shown in Fig. 1. The methodology section has been divided into two main phases: qualitative and quantitative.

3.1. Data collection of qualitative phase

This research is divided into two main parts. In the qualitative part, identification of the main indicators of customer churn by conducting interviews with insurance industry experts (managers, branch heads, technical and claims experts, and top representatives) has been made. The experts need to have experience in the insurance industry, and they should be aware of the structure of the insurance industry (e.g., managers; top representatives; superior organizational or individual insurees; consultants; top marketers; marketing professors from the insurance industry). As a result, experts' qualifications (interviewees) are defined. To participate in the survey, the experts should have at least 10 years of experience in the insurance industry. The data obtained from the interviews is collected and documented. Using the coding method, the indicators with the highest repetition are identified. Table 2 shows the summary and extraction of the interview indicators.

To examine the indicators more closely and determine the weights of each of them, in this part of the research, with the help of experts' opinions, validation of the selected indicators is performed. To perform paired comparisons and determine the degree of importance or weight of each indicator, an AHP questionnaire is designed and distributed among 20 other experts in the insurance industry.

3.1.1. Statistical population and sample in the qualitative phase

The statistical population is all managers of the insurance industry, heads of branches, deputies, top representatives, top insurers, and all clients of the insurance industry. According to the report of the Planning and Development Office of the Central Insurance Planning and Development Office of Iran, Iran Insurance Joint Stock Company has about 40 % of the insurance companies' share of the insurance premiums' production in the market. As a result, it can properly represent the industry, so that the managers of the insurance industry, heads of branches,

Table 1
Summarizing literature review and gap of research.

Author	Year	Country	Proposed method	Contribution of work	Software/data	Strategies	Industry
Youngjung, 2023	2023	Korea	Machine learning algorithm	Analyze the customer behavior information of actual water purifier rental company, where customer churn occurs very frequently, and to develop and verify the churn prediction model	Constructing a LABEL table, an analysis data mart, AWS, GCP-based customer data analysis platform, EDA	This study identified and calculated the influence of key variables on individual customer churn to enable a businessperson (rental care customer management staff) to carry out customer-tailored marketing to address the cause of the churn	Water purifier rental company
Leiria et al.	2022	Portugal	Binary logistic regressions	1. An improved understanding of the concept of customer loyalty; 2. The identification of the factors that insurance companies must manage to reduce product cancellation; 3. The importance of insurance companies to change their priorities from the acquisition to the retention of customers	IBM SPSS software	Determining internal and external practices of insurance companies improve their performance regarding product cancellation	Motor insurance industry
Chen et al.	2022	China	Cox model with variable penalties (Lasso, SCAD, and MCP)	Identifying the client churning and making comparisons among the penalties	Dataset from insurance company in China	Identifying clients who are prone to churn and reducing the maintenance cost of client management	Vehicle insurance
Robert Matthijs Verschuren	2022	Netherland	Extreme Gradient Boosting, XGBoost, multiple imputation	Employing a demand-based method to account for their indirect effect on a customer's willingness to stay	Automobile insurance portfolio from a large Dutch insurer	Presenting a causal inference framework for measuring customer price sensitivities and deducing (multi-period) optimal premium renewal offers	Automobile Insurance
Carolina Bellani	2019	Portugal	Logit models, random forests and artificial neural networks	Giving insights about the churn phenomenon and quantifies the churn risk for the considered policies	Client and conductor data	It can give a quantitative measure of the churn propensity, but it can also provide an understanding as to why this churn propensity occurs	Vehicle Insurance
Amin et al. (2019)	2019	Multinational	Presenting a novel CCP model based on distance factors to efficiently predict customer churns and estimate the level of certainty of the classifier's decision in each TCI dataset	Two main contributions to the existing literature: (i) introduced a novel approach for CCP in TCI based on distance factor, and (ii) revealed the effects of the distance factor in different distance zones (upper and lower zones) to estimate the expected certainty of the classifier decision.	MATLAB	Addressing the voluntary customer churns due to difficulty in predicting this type of customer churn while it is easier to filter out the involuntary customer churn by simple queries	Telecommunication
Jing et al. (2018)	2018	China	Rough set and BP neural network, Adam algorithm	The method proposed in this work is more robust and stable in dealing with a large number of dynamic data streams	VPRS-Adam-BP processing fluctuating data flow	BP neural network is applied to customer churn prediction, and variable precision rough set is introduced to preprocess the data, which removes incomplete information and reduces redundant attributes. Then, a feature extraction framework of variable precision rough set is established	Logistics industry
Tsai and Yu-Hsin, 2009	2009	USA	Hybrid data mining techniques by neural networks	ANN + ANN hybrid model significantly performs better than the SOM + ANN hybrid model and ANN baseline model.	Model evaluation by fuzzy testing data (FTD)	Providing better prediction accuracy and lower Type I errors, which is a better model for churn prediction	American telecom companies
Current research	*2023*	*Iran*	*Neural network and regression and AHP*	*A hybrid qualitative and quantitative model for strategizing and retaining organizational insurers.*	*MATLAB SPSS*	*Strategies to retain organizational insurers*	*Insurance industry*

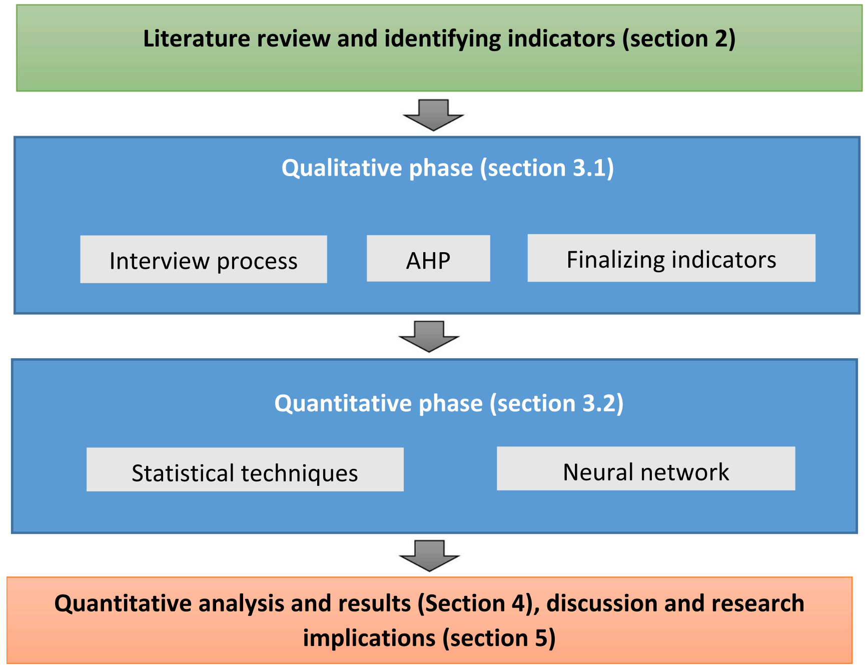

Fig. 1. Research framework.

technical and claims assistants, top representatives, and top insurers of Iran Insurance Company are interviewed. According to the population of respondents, samples are chosen by stratified sampling techniques and then judgmental sampling methods.

3.2. Data collection of quantitative phase

In this section, a hybrid model is designed using statistical techniques so that it can be used to predict insurees' churn while also classifying them. Several machine learning algorithms are commonly used to predict churn in different environments. The most widely used ones are logistic regression, decision tree, support vector machine and neural networks (Milosevic et al., 2017).

Several studies have proven that machine learning (ML) and artificial intelligence (AI)-driven analytics can extract data from customer relationship management (CRM) systems for the purpose of CCP and identify the root cause of churn. ML can help managers make informed and targeted decisions to retain customer churns (Amin et al., 2023). Some of the studies built churn prediction models based on the dynamic feature (Xing et al., 2022).

A neural network is a type of machine learning based on the model of the human nervous system using a series of mathematical and physical processes. Neural networks can be designed to perform tasks such as decision-making. Since being introduced in the 1950s, they've been used by IT professionals in a myriad of fields with great success. The most obvious applications of this technology are seen in classifiers and pattern recognition software. However, neural networks have also been successfully applied to predicting customer trends, highlighting newsworthy content, and even pricing models. Neural networks are beneficial for predictive analytics because they can synthesize complex and abstract relationships in your data, leading to greater predictive accuracy than some other methods. Neural networks are also very good at anticipating future trends and patterns that have not been observed in the past (Solomon, 2022). The research method in this research is as follows:

- Selecting 120 companies (insurees)
- Collecting data related to the eight above-mentioned indicators by checking the documents available at the insurance company.
- Using the neural network technique to analyze the data based on the following format (Table 2)
- Finally, clustering using K-Means and SPSS software.

The steps for the neural network model are shown in Fig. 2.

The process of grouping a set of data and putting them into classes of similar samples is called clustering. A cluster is a set of data that is like other data in the same cluster but different from the examples of other clusters. In the world of business, clustering helps marketers find distinct and different groups among their customers or identify customers based on purchasing patterns. One of the most important clustering techniques is the K-Means method, which is based on data analysis. This method is mostly used for big data. In this method, the number of clusters is assumed to be known, and the Euclidean distance method is used to determine the distance between two respondents. In this method, the initial data centers are selected first. Then, each of the repeated observations of the groups is added to that cluster based on the method of the closest Euclidean distance to the cluster mean. Also, the optimal number of clusters is determined. Therefore, the cluster centers change during each stage of development. This process continues until the average of the clusters does not change more than a certain value or until we reach the repetition limit.

3.2.1. Statistical population and sample in the quantitative phase

The statistical population in the quantitative phase is organizational insurees which is about 1250 companies. 120 samples are selected for analysis. In this research, data related to 120 insurance companies in different industries (namely clothing, metal, plastic, wood, etc.) are collected from documentation and records available in the database of Iran Insurance Company.

3.2.2. Data analysis method in the quantitative phase

In this phase, a neural network is used to validate the identified indicators, and K-means is used to cluster the policyholders.

Table 2
Summary of indicators obtained from interviews.

Indicator	1st interviewee	2nd interviewee	3rd interviewee	4th interviewee	5th interviewee	6th interviewee	7th interviewee	8th interviewee	9th interviewee	10th interviewee	11th interviewee	12th interviewee	13th interviewee	14th interviewee	15th interviewee	Total
Insurance premium amount	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓	✓	13
Dissatisfaction level				✓	✓				✓						✓	4
The number of insurance policies	✓	✓	✓	✓	✓		✓		✓		✓		✓		✓	10
Discount amount		✓			✓		✓	✓	✓	✓		✓				7
Number of complaints		✓		✓											✓	3
History of the policyholder						✓	✓	✓								3
Availability			✓						✓		✓	✓				4
Number of installments		✓	✓	✓	✓	✓	✓			✓				✓	✓	9
Speed of processing complaints			✓	✓		✓	✓		✓				✓	✓		7
Speed of insurance policy issuance							✓	✓	✓						✓	4
The way the staff behaves	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓				10
Amount of damages paid	✓		✓	✓	✓	✓		✓	✓		✓		✓	✓	✓	11
Number of covers		✓			✓			✓								3
Advertisements of insurers								✓							✓	2
Competitors' strategy		✓	✓		✓	✓	✓		✓	✓		✓				8
Geographical location					✓			✓		✓						3
Nature of the policyholder	✓	✓		✓		✓	✓									5
Extra service	✓															1
Honesty of the insurer	✓		✓				✓									3
Keeping the promises/ commitments	✓						✓						✓			3
Damage declaration frequency	✓	✓			✓			✓			✓		✓	✓	✓	8
Bad economic conditions		✓	✓				✓		✓							4
Speed of insurance policy delivery			✓					✓								2
Method of receiving insurance premiums		✓	✓			✓	✓			✓	✓	✓		✓		8
Lack of liquidity		✓				✓			✓	✓						4

(*continued on next page*)

Table 2 (*continued*)

Indicator	1st interviewee	2nd interviewee	3rd interviewee	4th interviewee	5th interviewee	6th interviewee	7th interviewee	8th interviewee	9th interviewee	10th interviewee	11th interviewee	12th interviewee	13th interviewee	14th interviewee	15th interviewee	Total
Type of insurance policy	✓	✓		✓	✓		✓	✓		✓					✓	8
Insured value		✓													✓	2
Tender holding		✓		✓		✓										3
Method of introduction			✓													1
Number of visits				✓											✓	2
Quality of service				✓						✓	✓		✓	✓		5
Advertising method					✓							✓				2
History of cooperation/ renewal		✓	✓	✓	✓	✓		✓		✓				✓	✓	9
The number of personnel in the insurance company					✓	✓										2
Number of branches/ representatives						✓			✓	✓						3
The physical aspect of the workplace							✓			✓						2
Speed of damage handling	✓						✓					✓		✓	✓	5
Franchise percentage		✓													✓	2
History of the insurer				✓											✓	2
Ethics of the marketer							✓				✓	✓				3
Risk taking power									✓							1
The credit of the insurer or insurance agent										✓	✓	✓				3
After-sales service											✓					1
public satisfaction											✓					1
Image of the brand											✓					1
Informing method												✓				1
Informing time												✓				1
Non-transparent contracts														✓		1

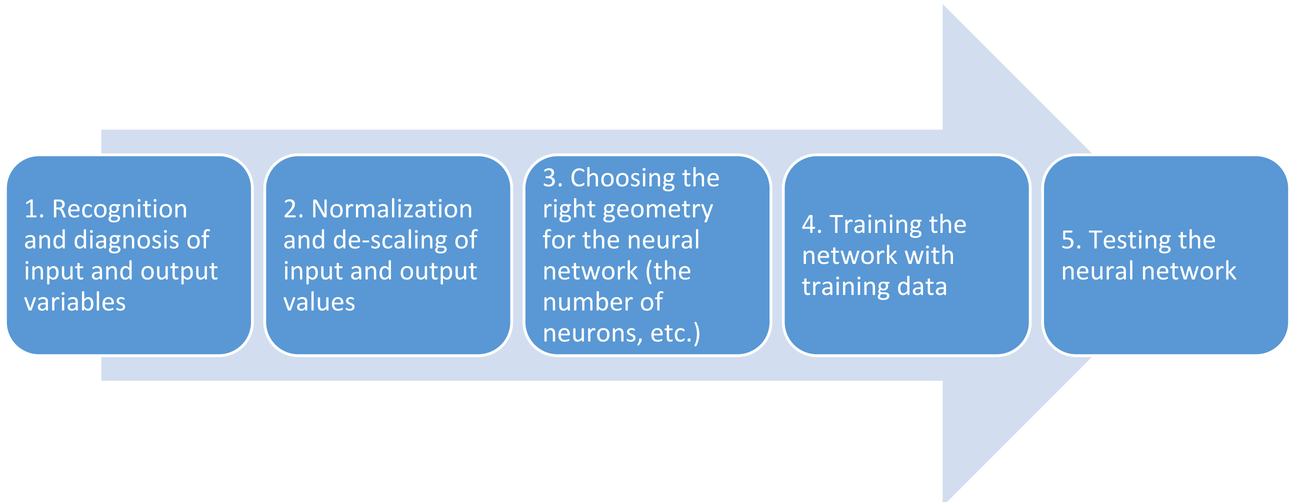

Fig. 2. The steps for the neural network model.

4. Results

4.1. Research results from the qualitative phase

First, the data from the interviews are collected. The interview indicators are shown in Table 2.

To examine the indicators more closely and determine the weights of each of them, the AHP technique is applied. To apply the AHP technique, a paired comparison questionnaire is designed. The questionnaire is given to 20 experts, apart from the 15 people who were interviewed earlier. These experts all have educational and experiential records related to the insurance industry. These questionnaires are analyzed using Expert Choice software, the output of which is according to Table 3 and Fig. 3.

After applying AHP techniques, all criteria are ranked based on ideal values, as shown in Fig. 3.

To select the final indicators, the following decision-making rules are adopted:

A. Examining the factors identified and extracted from the research literature.
B. Examining and extracting indicators from interviews conducted with experts.
C. Specifying the number of repetitions of indicators in the research literature
D. Specifying the number of repetitions of indicators in interviews
E. Determining the number 8 as the basis for selecting the index (at least 8 repetitions in the literature and interview)
F. Considering the existence of quantitative data regarding the selected indicators

Table 3
The result of AHP technique.

Indicators	Ideals	Normals	Raw
Tender holding	0.148784	0.37668	0.037668
Number of premium installments	0.408441	0.103406	0.103406
Number of insurance policies	0.524981	0.132911	0.132911
Number of claim announcements	0.201614	0.051043	0.051043
The length of the relationship	0.172228	0.043603	0.043603
History of the insurer	0.042263	0.010700	0.010700
Nature of the policyholder	0.063721	0.016132	0.016132
Insurance premium amount	0.1000000	0.253172	0.253172
The final result of claims	0.750795	0.190081	0.190081
Insurance premiums' payment method	0.401225	0.101579	0.101579
Type of insurance policy	0.235826	0.059705	0.059705

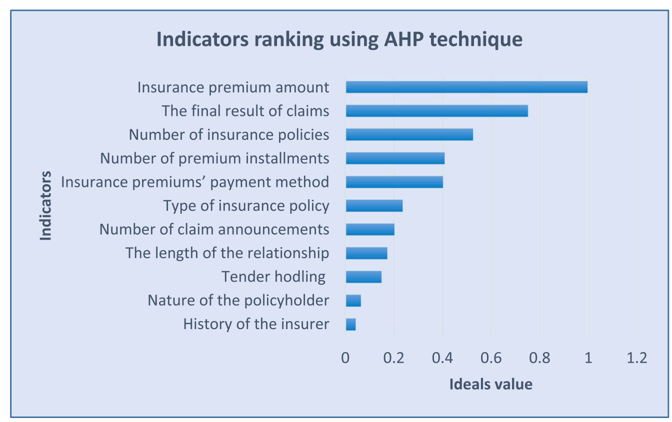

Fig. 3. The result of AHP technique.

G. The results of the AHP technique

According to the above decision rules, eight indicators are finalized, which are:

1. Insurance premium amount
2. Number of insurance policies
3. Number of premium installments
4. Number of claim announcements
5. The length of the relationship or the number of times the insurance policy is renewed.
6. The final result of claims (amount of claims)
7. Insurance premiums' payment method
8. Type of insurance policy

Three types of insurance policies are chosen: fire insurance, liability insurance, and health insurance. These policies have experienced a large amount of churn based on the data within 5 years; meanwhile, they include a significant part of the portfolio of insurance companies. The database of Iran's insurance industry reflects a significant rate of churn among organizational insurees in 2022, particularly in fire insurance, liability insurance, and health insurance. According to Iran Central Insurance's annual report, the number of liability insurance policies decreased to 4100 in 2022 from 5200 in 2021 in only one year. Fire insurance policies dropped to 11,543 from 12,750 in the same period. This happened for health insurance as well. The number of policies dropped to 632 from 734 (Central Insurance of Iran, 2022).

4.2. Research results from the quantitative phase

4.2.1. Results from the neural network

The finalized eight indicators are coded with the neural network technique, and an analytical model resulting from the application of the neural network is obtained, which shows how the eight factors can best justify the reason for organizational insurees' churn. Following the implementation of the relevant codes, the results are shown in Fig. 4.

The neural network training process is shown in Fig. 5. The number of repeats (APACs) is set at 12. The optimal state occurred in Apac 9, and the error in this case is 0.15085. In fact, point 0.15058 is the best case with the least error (Fig. 6).

The model and the regression value of the predicted neural network are finally $R = 0.74561$ in Fig. 5.

R ranges between 0 and 1 ($0 \leq R \leq 1$). The closer it is to 1, it indicates that the indicators are better able to predict and explain the churn. The value of regression obtained from the neural network method is $R = 0.74561$. Since it is >0.5, it means that the model is significantly acceptable. The resulting analytical model is illustrated in Fig. 7.

The most important factors influencing the churn of insurees in Iran's insurance industry could be recognized, and the indicators are (Fig. 8):

1. Number of insurance policies that an insurer buys from the insurance company.
2. The types of insurance policies purchased (in this study, fire insurance, liability insurance, and health insurance are examined)
3. The amount of insurance premiums (in this study, fire insurance, liability insurance, and health insurance are examined)
4. The number of claims declared (in this study, fire insurance, liability, and health insurance are examined)
5. The average number of installments of the insurance policy (in this study, this number is herewith obtained: the total number of installments divided by the number of types of insurance policies purchased)
6. How to pay the premium (in this study, i.e., cash payment or payment by check)
7. The length of the relationship with the insurer, which refers to the number of times the insurer has renewed the insurance policy with the insurer.
8. The result of the claim cases, which means the general satisfaction of the insurer with the claim of loss and the amount that he receives for loss,.

These eight factors influencing the churn of corporate insurees are analyzed as final indicators. The result of the neural network technique (calculating $R = 0.73$) indicates that the identified factors can best explain the reason for the churn of corporate insurees.

4.2.2. Result from K-means method for clustering

The K-means method provides the opportunity to predict customer churn. The second research question is related to designing a model for predicting the churn of organizational insurance policyholders. By using

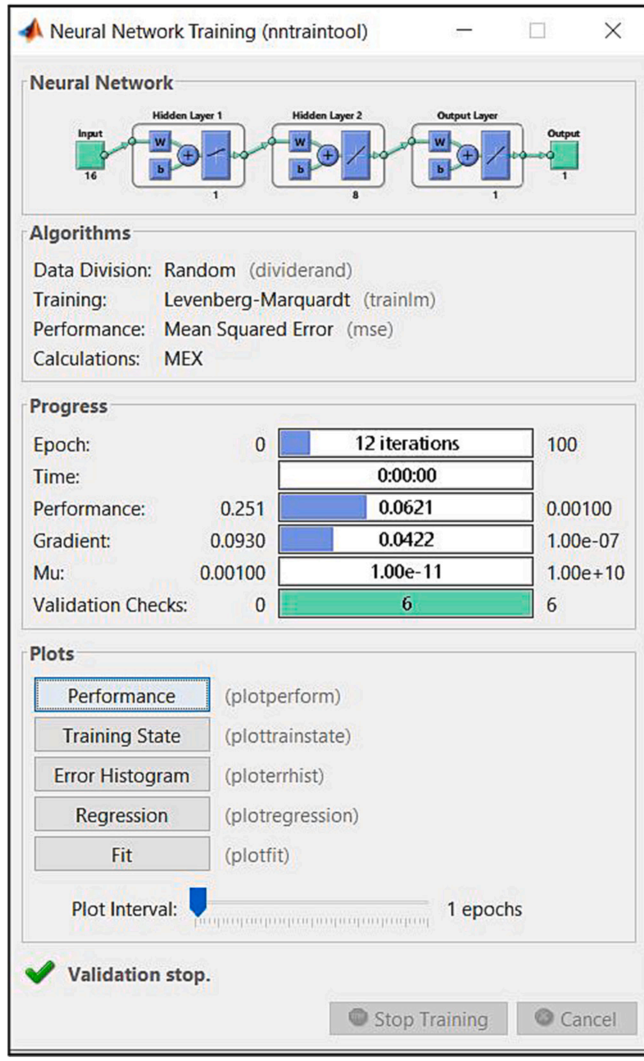

Fig. 5. The neural network training process.

the K-Means technique, organizational insurees could be divided into two categories: churned insurees and retained insurees. This model is based on real data. One of its main and practical features is that if a new policyholder, apart from the examined samples, is examined in this model based on the required criteria, the rate of churn for the coming year can be predicted. This means that we can identify if the policyholder will be maintained or churned.

K-means clustering is an extensively used technique for data cluster analysis. Clustering is one of the data mining methods that partition large-sized data into subgroups according to their similarities. K-means clustering algorithm works well in spherical or convex data distribution of large-sized data sets. Most of the algorithms based on K-means have generally been interested in an initial cluster centers selection or cluster

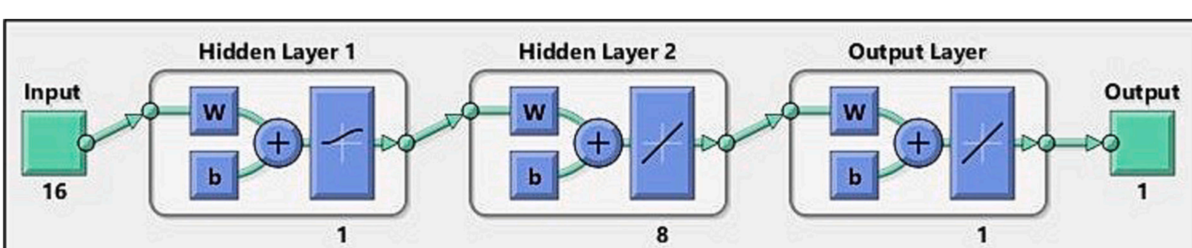

Fig. 4. Neural network image related to model design in MATLAB software.

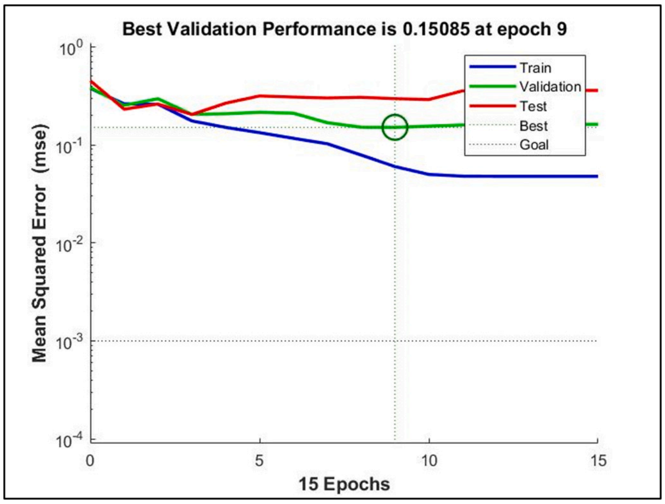

Fig. 6. Performance of perceptron artificial neural network in optimal condition.

distribution (Ay et al., 2023).

In this clustering, all 16 features (inputs) are considered for policyholders. Table 4 presents the variables used in K-Means.

Table 5 is the table of initial cluster centers and represents the initial average of each index within each cluster.

Table 6 shows the progress of the clustering in each step. In the initial iterations, the cluster centers are slightly changed to reach stability (zero). The clustering method is done via 5 repetitions.

Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is 0.000. The current iteration is 5. The minimum distance between initial centers is 3700.178.

Table 7 specifies in which cluster each company/customer/policyholder falls into. For example, customer 1 is in the first cluster and customer 2 is in the second cluster. The fourth column also specifies the distance of each customer from the center of the cluster. If this distance is high, it indicates that the cluster is not a good representative for its clustering. Among 120 samples, 109 one has valid data.

Table 8 shows the Euclidean distances between the final cluster centers. According to the displayed number, the distance between two clusters is large.

Table 9 shows which variable plays the most important role in clustering (note the F value). According to the displayed numbers, the variable "health care insurance price" has the most effective role in separating the clusters at an error level of <0.01, and on the contrary, the variable "liability insurance amount" has the least important role at an error level of <0.978 in the separation of clusters.

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal. Table 10 shows the number of customers in each cluster; 83 customers are in cluster 1 and 26 customers are in cluster 2.

Based on the above points and the outputs of the model, the variables can be illustrated in the order of effect in Table 11. According to the results, the variable "health care insurance price" has the most effective role in separating the clusters at an error level of <0.01, and on the

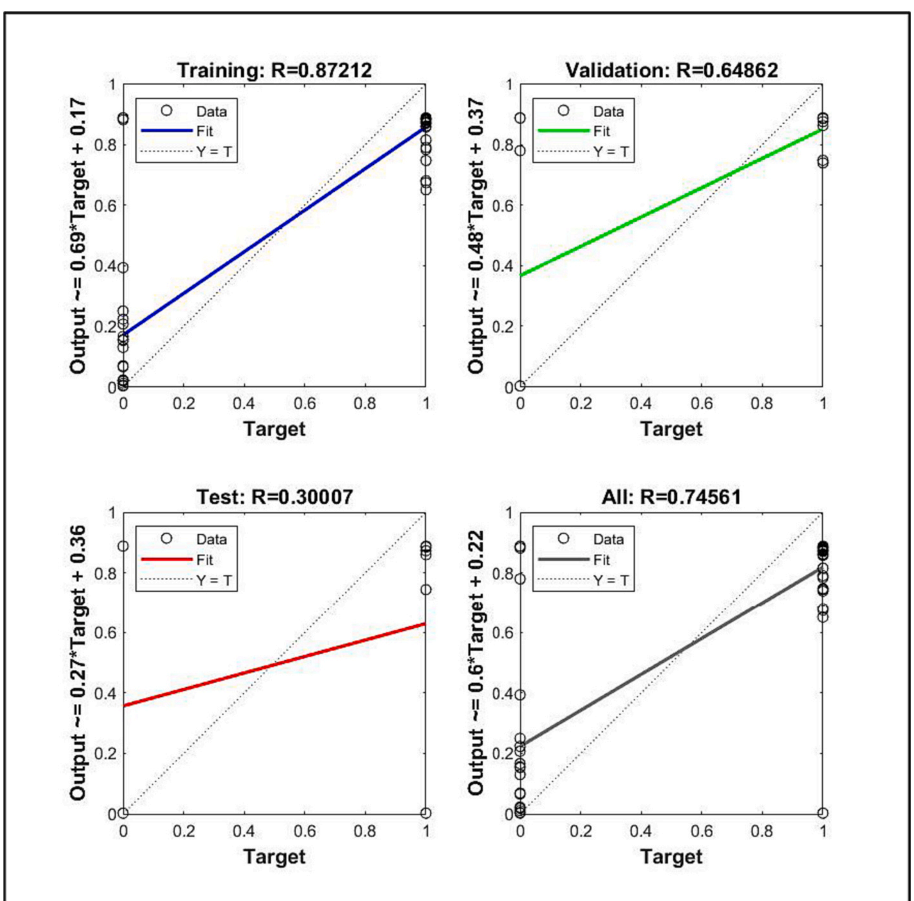

Fig. 7. Regression model.

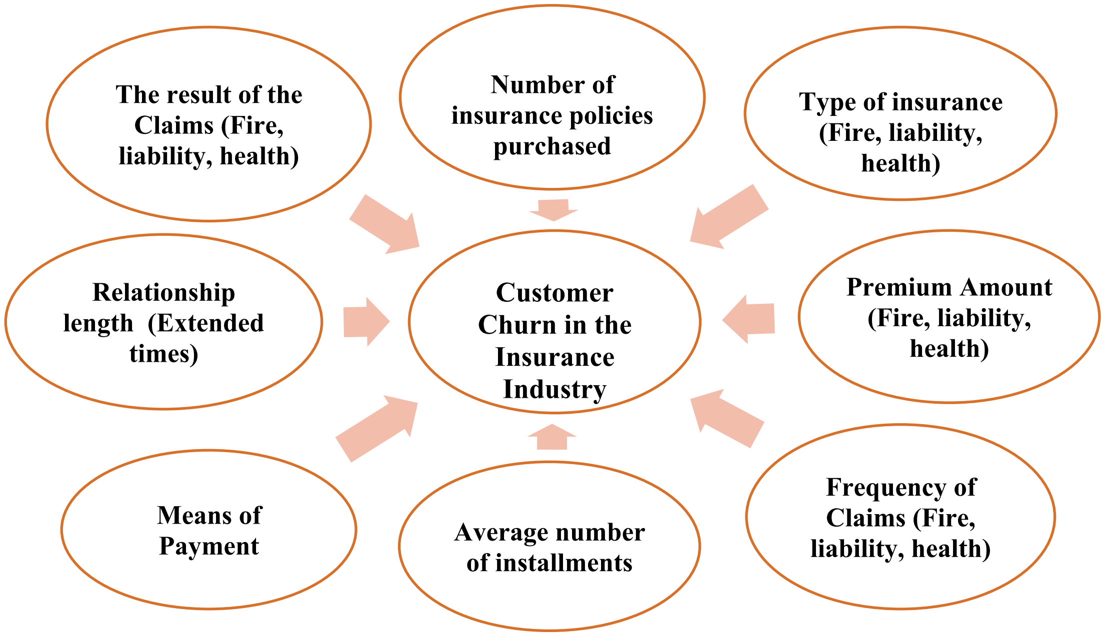

Fig. 8. Analytical research model.

Table 4
K-means variables.

No.	Variable (inputs)	Variable description
1	Number of insurance policy	The number of insurance policies purchased
2	Bime Atashsouzi	Type of insurance (fire)
3	Bime Masouliat	Type of insurance (liability)
4	Bime Khadamat Darmani	Type of insurance (health)
5	IPF	Fire insurance premium amount
6	IPM	Liability insurance premium amount
7	IPD	Health insurance premium amount
8	NDF	The frequency of fire claims
9	NDM	Frequency of liability claims
10	NDD	Frequency of health claims
11	Average of installments	Average number of installments
12	Type of payment	Payment method for insurance premiums
13	Number of extend	Length of relationship (renewal times)
14	KhesaratF	Final result of fire claims
15	KhesaratM	Final result of liability claims
16	KhesaratD	Final result of health claims

Table 5
Initial cluster centers.

Variable description	Cluster	
	1	2
The number of insurance policies purchased	2	1
Type of insurance (fire)	1	0
Type of insurance (liability)	0	0
Type of insurance (health)	1	1
Fire insurance premium amount	36	0
Liability insurance premium amount	1200	0
Health insurance premium amount	0	3500
The frequency of fire claims	0	0
Frequency of liability claims	0	0
Frequency of health claims	9	11
Average number of installments	8	12
Payment method for insurance premiums	1	1
Length of relationship (renewal times)	1	2
Final result of fire claims	0.00	0.00
Final result of liability claims	0.00	0.00
Final result of health claims	0.17	1.00

Table 6
The progress of the clustering process at each stage.

Iteration history		
Iteration	Change in cluster centers	
	1	2
1	1191.818	1390.971
2	97.989	266.814
3	44.426	108.087
4	9.724	28.659
5	0.000	0.000

Table 7
Status of some insurees in each cluster.

Case number	Group name	Cluster	Distance
1	Metal industry	1	588.730
2	Wood industry	2	517.927
3	Plastic industry	2	80.135
.			
.			
.			
109	Petrochemical industry	1	1152.340

Table 8
Euclidean distances between final cluster centers.

Distance between final cluster centers		
Cluster	1	2
1		1522.304
2	1522.304	

contrary, the variable "liability insurance amount" has the least important role at an error level of <0.978.

According to the obtained results, the insurance company should adopt a specific strategy for those organizational policyholders who are currently maintained but are on the edge and are most likely to churn in the next year. They probably will not renew their policies for the next year. Therefore, the insurance company should invite them or pay

Table 9
Comparing the effectiveness of variables, ANOVA.

Variable description		Cluster		Error		F	Sig.
		Mean square	df	Mean square	df		
The number of insurance policies purchased	Number of insurance policy	5.304	1	0.526	107	10.089	0.002
Type of insurance (Fire)	Bime Atashsouzi	1.141	1	0.153	107	7.434	0.007
Type of insurance (Liability)	Bime Masouliat	0.237	1	0.251	107	0.942	0.334
Type of insurance (Health)	Bime Khadamat Darmani	8.633	1	0.172	107	50.116	0.000
Fire insurance premium amount	IPF	73,322.409	1	24,761.470	107	2.961	0.088
Liability insurance premium amount	IPM	21.580	1	27,689.598	107	0.001	0.978
Health insurance premium amount	IPD	45,805,850.01	1	152,785.205	107	299.806	0.000
The frequency of fire claims	NDF	0.048	1	0.753	107	0.063	0.802
Frequency of liability claims	NDM	0.197	1	0.901	107	0.218	0.641
Frequency of health claims	NDD	1169.373	1	39.648	107	29.494	0.000
Average number of installments	Average of installments	93.871	1	4.631	107	20.270	0.000
Payment method for insurance premiums	Type of payment	0.123	1	0.189	107	0.651	0.421
Length of relationship (renewal times)	Number of extend	0.888	1	2.634	107	0.337	0.563
Final result of fire claims	KhesaratF	0.003	1	0.144	107	0.020	0.888
Final result of liability claims	KhesaratM	0.266	1	0.131	107	2.027	0.157
Final result of health claims	KhesaratD	2.901	1	0.087	107	33.382	0.000

Table 10
Number of cases in each cluster.

Cluster	1	83.000
	2	26.000
Valid		109.000
Missing		0.000

Table 11
Order of the effectiveness of model variables.

Row	Description of variables
1	Health Insurance Premium
2	Type of insurance (Health Insurance)
3	The result of Health Insurance Claims
4	Frequency of claims in Health Insurance
5	Average number of installments
6	Number of insurance policies purchased
7	Type of insurance (fire)
8	Fire Insurance Premium
9	The result of Liability Claim
10	Type of insurance (Liability)
11	Means of Payment
12	Relationship Length (Extended Frequency)
13	Frequently of Claims in Liability Insurance
14	Frequency of Claims in Fire Insurance
15	The result of Fire Insurance Claims
16	Liability Insurance Premium

special attention to them in its customer relationship management policies. The company can pay their claims within a week, give a special discount on their fire insurance premiums, or increase the number of installments of their insurance policies. Since some policyholders are more likely to churn in the next year, the insurance company is advised to pay special attention to them. By slightly raising the level of service to them, the insurance company can easily retain these organizational insurers and increase its retention rate. For the first group, it is better to calculate their profitability over the past few years. If they are profitable, the CRM is suggested to hold urgent and private meetings with them, investigate the reasons for churn, and adopt strategies to retain them (very high discounts, multiple installments, quick payment of their claim cases, etc.). But if the analysis shows that they are unprofitable insurers, the insurance company can revise its 4P strategy in such a way that, for example, it increases insurance premiums, implies more control over the amount of payment for their claims, or tries to issue an insurance policy for them in cash.

For the second group, after calculating their profitability level, if they are profitable, then the 4P policies should be slightly revised. Because of slight changes in the level of service provided to them, it is easy to retain them in the insurer's portfolio and increase the retention rate. But if they are unprofitable, then they should be treated with more caution. For example, to act more cautiously in the discounts or in the installments of their insurance policies, try to collect the installments in cash and at shorter time intervals.

For the third group, after calculating their profitability level, if they are profitable, special attention should be paid to them because they have a higher potential to become loyal insurers. They must be treated as VIPs. Special discounts should be given to them, their claim cases should be handled out of turn, they should determine the terms of payment of insurance premiums themselves, etc. However, if the insurers are unprofitable, they should revise the 4P strategies. Since they have the potential to become loyal insurers while being unprofitable, they need to order the issuance of other insurance policies with the same insurance company or introduce the insurance company to other friends and acquaintances. The insurance company should be more careful when giving premiums to this group and avoid bold discounts. While maintaining the current level of service, attempts should be made to influence these insurers and sell them more insurance policies.

For the fourth group, after calculating their profitability level, if they are profitable, they should be treated as VIPs with maximum attention. They are very vulnerable and relatively sensitive to competitors' strategies and have a high potential to be attracted to rival insurance companies. Therefore, the level of service to them should be greatly increased. Discounts should be given to them, their damage cases should be investigated and paid attention to, and they should be appreciated or adopt other 4P strategies. But if they are unprofitable, they should be treated like the third group.

5. Conclusion and implication of research

Identifying and selecting key indicators to predict the churn of institutional insurees is one of the outstanding challenges facing insurance companies. In this study, to help insurance companies formulate effective marketing strategies for policyholders' churn, the key indicators are first identified in the research literature, including 48 items. Then, to select the most effective indicators, the ones with at least 8 repetitions are chosen. To validate these indicators and determine their weights, the AHP hierarchical analysis method is used. In the qualitative part of the research, eight indicators are finalized (answers to the first research question). In other words, the result of the study in the qualitative part proves that the following criteria are the main reasons for policyholders' churn in the Iranian insurance industry: number of insurance policies purchased, type of insurance, premium amount, frequency of claims, average number of installments, means of payment, relationship length, and result of the claims. Afterwards, the neural network method is applied to validate the indicators. Achieving $R = 0.74561$ indicates that the obtained indicators can predict 75 % of the behavior of corporate insurance policyholders. Afterwards, to help insurance companies identify which organizational insuree will churn or is maintained in the company's portfolio within the coming year, the clustering approach was employed. The clustering results determined the number of churned or retained organizational insurees for Iran Insurance Company. Finally, to identify policyholders who will churn, the K-means technique is applied. This means that it is possible to design a model, with the help of statistical and mathematical techniques, to categorize insurees in Iran's insurance industry and predict their churn (answer to the second research question). The output of the model leads us to adopt proper and practical marketing strategies and customer relationship management policies. Some of these practical findings are as follows.

Considering that healthcare insurance policy is the most effective factor affecting churn among institutional insurers, it is therefore recommended to collect the claim files at short intervals and proceed to pay insurance damages within a maximum period of 15 days. Since the average index of the number of installments of insurance policies shows that it has the greatest impact on the churn of institutional policyholders, it is recommended to consider the number of installments of insurance policies for at least 10 months to maintain and satisfy the policyholders. The number or variety of insurance policies that are purchased by the corporate insurer during a one-year period from the insurance company also has a great impact on the policyholders' churn. Examining other indicators in relation to each of the three insurance policies (fire insurance, liability insurance, and healthcare) shows that, especially in relation to liability insurance, the amount of insurance premium has a very small effect and the amount of fire insurance has a medium impact. As a result, the variety of purchased insurance policies can somehow reduce the churn of policyholders. Because dissatisfaction in the field of medical insurance policies can be somehow compensated by satisfaction with other insurance policies,.

The way insurance premiums are paid also has a medium effect on the churn of institutional insurers. Since most insurers refuse to pay by check, it is suggested that, as far as possible, the insurance premiums be received in cash at intervals of the specified installment dates and avoid pressure to receive checks. On the other hand, practical experience shows that policyholders who voluntarily pay checks are prudent policyholders and intend to remain in the insurer's portfolio and continue to cooperate with the same insurer.

It is recommended to design different marketing strategies according to the four groups of insurers that have been predicted and identified by the model. The first group is policyholders who will churn, and the possibility of their return is very weak. The second group is made up of policyholders who have turned away, but the probability of their return is very high. The third group is made up of policyholders who have been retained and will renew their insurance policies in the coming years. The fourth group: insurers who have been retained but are likely to leave the portfolio composition in the next year or have high potential to churn.

Despite this, in the analysis of the loss of institutional insurers in the insurance industry, it can be stated that since an insurance company—the customer—has several insurance policies from insurance companies, it is difficult to accurately analyze and diagnose the loss because it is possible that an insured who has three insurance policies from insurance companies and does not renew one of his insurances in the following year cannot accurately acknowledge that the loss has occurred. In this research, the loss of the policyholder occurs when at least one policy has not been renewed in the following year. In the insurance industry, a policyholder or insuree has several insurance policies, and it is very difficult to accurately analyze and diagnose the churn because a policyholder might have three insurance policies and not renew one of the policies in the following year. This means that churn cannot be precisely recognized. In this research, the churn occurs when at least one insurance policy has not been renewed in the following year.

For further research, it is suggested to test the prediction model with a larger number of samples. Moreover, it would be worthwhile to investigate the influential factors on the churn of individual insurers and the prediction models related to it. And finally, it is suggested to model and describe the influential factors in the churn of organizational insurers with the help of other forecasting techniques.

Ethical approval

The authors declare no conflict of interest.

Consent to participate

The authors declare that they agree with the participation of the journal.

Consent to publish

The authors declare that they agree with the publication of this paper in this journal.

Funding

There is no funding reported in this paper.

CRediT authorship contribution statement

Samaneh Soltani: Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Adeleh Eisazadeh Saravani:** Software, Resources, Project administration, Data curation. **Kambiz Shahroodi:** Writing – review & editing, Writing – original draft, Supervision, Software, Conceptualization. **Soroush Avakh Darestani:** Writing – review & editing, Validation, Supervision, Project administration.

Declaration of competing interest

Not applicable.

Data availability

The authors declare that the data can be presented upon the requested of the readers.

Appendix 1. Interview questions

The interview protocol is consisted of the following open questions:

1. Please introduce yourself briefly (name and surname/position/work experience/company etc.).
2. In your company/in your opinion, what is the meaning of churn of a policyholder?
3. In your opinion, when a policyholder churn happens?
4. In your opinion, what are the factors that cause a corporate insuree to churn?

- What are the individual or organizational factors?
- What factors are related to the insurance agent?
- What factors are related to the insurance marketer?
- What factors are related to insurer services?
- What factors are related to the insurance policy?
- What factors are related to the environment?

5. Can you specify the cause-and-effect relationships between some factors?
6. In your opinion, how can these factors be ranked in terms of importance?

Some sample answers:

Interviewee 4	"The most important factor is honesty, and after that, the way employees treat customers. Then it seems that building trust and speed of handling claims are important. Afterwards, the amount of insurance premium and then insurance compensation are important. The number of issued insurance policies and the duration of cooperation are the next priorities. Competitors' strategies, economic conditions, the way and speed of insurance policy delivery, insurance premium payment methods and terms are next in importance."
Interviewee 7	"The amount of insurance compensation and the amount of insurance premium are the first and the second priorities. Then, the number of installments, the length of cooperation, the amount of liquidity, the nature of the policy holder, economic factors, negative advertisements of competitors, the flexibility of competitors, the number and type of purchased insurance policies can be mentioned."
Interviewee 11	"The most important indicators are, respectively, treating customers with respect, the amount of insurance premium, the amount of compensation, the number and type of insurance policies purchased, the method of installments, the satisfaction of the policyholder, the nature of the insurer (whether it is public or private), the duration of cooperation, the history of the insurer, and the number of insurance objections, the frequency of visiting the company and…"

References

Amin, A., Al-Obeidatb, F., Shahab, B., Adnana, A., Loo, J., Anwar, S., 2018. Customer churn prediction in telecommunication industry using data certainty. J. Bus. Res. 1–12.

Amin, et al., 2019. Customer churn prediction in telecommunication industry using data certainty. J. Bus. Res. 94, 290–301. January 2019.

Amin, Adnan, Adnan, Awais, Anwar, Sajid, 2023. An adaptive learning approach for customer churn prediction in the telecommunication industry using evolutionary computation and Naïve Bayes. Appl. Soft Comput. 137, 1–16. https://doi.org/10.1016/j.asoc.2023.110103.

Ay, Merhad, Özbakır, Lale, Kulluk, Sinem, Gülmez, Burak, Öztürk, Güney, Özer, Sertay, 2023. FC-Kmeans: fixed-centered K-means algorithm. Expert Syst. Appl. 211, 118656. ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2022.118656. https://www.sciencedirect.com/science/article/pii/S0957417422016979.

Bellani, C., 2019. Predictive Churn Models in Vehicle Insurance (Master thesis). Universidade Nova de Lisboa. https://core.ac.uk/download/pdf/303770784.pdf.

Bolancé, C., Guillen, M., Padilla-Barreto, A.E., 2016. Predicting probability of customer churn in insurance. In: León, R., Muñoz-Torres, M., Moneva, J. (Eds.), Modeling and Simulation in Engineering, Economics and Management. MS 2016. Lecture Notes in Business Information Processing, vol 254. Springer, Cham.

Bravante, Joshua James A., Robielos, Rex Aurelius C., 2022. Game over: an application of customer churn prediction using survival analysis modelling in automobile insurance. In: Proceedings of the International Conference on Industrial Engineering and Operations Management Istanbul, Turkey, March 7–10, 2022.

Castro, Laura Sofia Sauthoff da Ponte, 2022. Customer Churn Prediction in Insurance: Modeling Renewal Price Elasticity of the Workers' Compensation Portfolio From Ocidental Seguros. Internship Report presented as the partial requirement for obtaining a master's degree in data science and advanced Analytics, May 2022. https://run.unl.pt/handle/10362/141571.

Central Insurance of Iran, 2022. Iran Central Insurance's annual report. www.centinsur.ir.

Chen, S.-H., 2016. The gamma CUSUM chart method for online customer churn prediction. Electron. Commerce Res. Appl. 17, 99–111.

Chen, Yan, Zhang, Lei, Zhao, Yulu, Xu, Bing, 2022. Implementation of penalized survival models in churn prediction of vehicle insurance. J. Bus. Res. 153 (C), 162–171. Elsevier.

De Caigny, A., Coussement, K., De Bock, W., K., 2018. A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. Eur. J. Oper. Res. 269, 760–772.

Günther, Clara-Cecilie, Tvete, Ingunn Fride, Aas, Kjersti, Sandnes, Geir Inge, Borgan, Ørnulf, 2014. Modelling and predicting customer churn from an insurance company. Scand. Actuar. J. 1, 58–71. https://doi.org/10.1080/03461238.2011.636502, 2014.

Huigevoort, C., 2015. Customer Churn Prediction for an Insurance Company (Master Thesis). Eindhoven University of Technology.

Jing, Gong, et al., 2018. Research on customer churn prediction method based on variable precision rough set and BP neural network. Adv. Intelligent Syst. Res. 161, 287–293.

Leiria, Manuel, Rebelo, Efigénio, de Matos, Nelson, 2022. Measuring the effectiveness of intermediary loyalty programmes in the motor insurance industry: loyal versus non-loyal customers. Eur. J. Manag. Bus. Econ. (Issue publication date: 3 May 2022).

Liu, Zhenkun, Jiang, Ping, De Bock, Koen W., Wang, Jianzhou, Zhang, Lifang, Niu, Xinsong, 2024. `Extreme gradient boosting trees with efficient Bayesian optimization for profit-driven customer churn prediction. Technol. Forecast. Soc. Chang. 198, 122945.

Matthijs Verschuren, Robert, 2022. Customer price sensitivities in competitive insurance markets. Expert Syst. Appl. 202, 117133, 2022. ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2022.117133. https://www.sciencedirect.com/science/article/pii/S0957417422005309.

Milosevic, M., Zivic, N., Andjelkovic, I., 2017. Early churn prediction with personalized targeting in mobile social games. Expert Syst. Appl.

Prabadevi, B., Shalini, R., Kavitha, B.R., 2023. Customer churning analysis using machine learning algorithms. Int. J. Intelligent Networks 4, 145–154. https://doi.org/10.1016/j.ijin.2023.05.005.

Reichheld, F.F., Sasser Jr, W.E., 1990. Zero Defections: Quality Comes to Services. Harv. Bus. Rev. Magazine Article (Accessed 1 November 2024).

Solomon, Martin, 2022. Predictive Analytics Using Neural Networks. https://www.linkedin.com/pulse/predictive-analytics-using-neural-networks-martin-solomon/?trackingId=Sey2cDLqR1K0T5ulbaE3UA%3D%3D.

Soltani Lifshagerd, S., Shahroodi, K., Chirani, E., 2021. Predicting customer churn in the insurance industry: identifying the influential factors. J. Invest. Knowl. 10 (39), 341–354. https://ijfma.srbiau.ac.ir/article_16891.html.

Tsai, Tsai Chih-Fong, Yu-Hsin, Lu, 2009. Customer churn prediction by hybrid neural networks. Expert Syst. Appl. 36, 12547–12553.

Usman-Hamza, Fatima E., Balogun, Abdullateef O., Nasiru, Salahdeen K., Capretz, Luiz Fernando, Mojeed, Hammed A., Salihu, Shakirat A., Akintola, Abimbola G., Mabayoje, Modinat A., Awotunde, Joseph B., 2024. Empirical analysis of tree-based classification models for customer churn prediction. Sci. Afr. 23, e02054.

Wagh, Sharmila K., Andhale, Aishwarya A., Wagh, Kishor S., Pansare, Jayshree R., Ambadekar, Sarita P., Gawande, S.H., 2024. Customer churn prediction in telecom sector using machine learning techniques. In: Results in Control and Optimization, 14, p. 100342.

Xing, Wu, Li, Pan, Zhao, Ming, Liu, Ying, Crespo, Rubén González, Herrera-Viedma, Enrique, 2022. Customer churn prediction for web browsers. Expert Syst. Appl. 209, 118177 https://doi.org/10.1016/j.eswa.2022.118177.

Youngjung, Suh, 2023. Machine learning based customer churn prediction in home appliance rental business. J. Big Data 10, 41. https://doi.org/10.1186/s40537-023-00721-8.