Attributes priority arrangement of cancellations in the life insurance using rough set

[a*]M. A. Elsafty, [b] H. M. Elkhawaga, [c] A.M.Kozae

[a*] Department of Mathematics and Statistics, Faculty of Science, Taif University, elsafty010@yahoo.co.uk

[b]Department of Statistics and Mathematics and Insurance, Tanta University, Egypt.

[b]Department of Investment and Finance, Taif University, Kingdom of Saudi Arabic.

hamedelkhwaga70@yahoo.com

[c] Department of Mathematics, Faculty of Science, Tanta University, Egypt, akozae55@yahoo.com

Abstract:

In this paper, we present a new method to rough set attributes priority arrangement of cancellations in the life insurance, the number of canceled documents on the rise plus they oscillatory also in rates of cancellations in constant fluctuation which requires searching in the problem of cancellations of these documents where, these cancellations affect negatively the results of the Egyptian life insurance operations of insurance companies. This paper also aims to identify the most important variables that affect the rates of cancellations and so as to reach solutions that lead to reducing the number of cancelled documents by using the Rough Model. The research also aims to use the Rough model to identify the determinants of cancellations in the life insurance of the Egyptian insurance market, so as to reach solutions that lead to reducing the number of cancelled policies and find solutions to them. Our findings proved that the most important factors that affect the document cancellation in the Egyptian life insurance policies are; policy premiums, insurance period and age of policyholders respectively.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AMS Subject Classification: **C020** Mathematical Methods, **Y100** Data: Tables and Charts

Keywords: Rough set, Life Insurance, cancellations, Lower and Upper Approximations, Core and Reduction, Accuracy.

Introduction

The theory of rough set has been successfully applied to diverse areas, such as pattern recognition, artificially intelligent, machine learning, knowledge acquisition, economy forecast, and data mining [3, 5, 18-19]. Pawlak [15-17] rough set model is constructed based on equivalence relations. These relations are studied by many investigators to be used in the complex decision tasks. In multiple criteria decision-making problems, there are preference structures between conditions and decisions [1, 6]**.** A reduct should be able to preserve the original classification power provided by the whole attribute set [2, 12]. This power may be interpreted by syntax properties and semantics properties for both positive and boundary rule sets. Instead, we need to consider multiple properties and multiple measures for evaluation. This paper addresses different criteria, such as confidence, coverage, generality, cost, and decision-criteria based on the decision-theoretic rough set models [8, 12, 22].

The classical rough set methods cannot detect the inconsistency related to the preference, such as the price, fuel amount, speed, etc., and these attributes involve preference information but they are not considered in the rough set [11, 13]. If we replace the indiscernibility relation with the dominance relation to restructure the rough set models, we can extend rough set models as methods to solve problems met to analyze preference with multiple attributes, and these hybrid models have not only the best quality of classical rough set models, but also the ability to deal with the possible inconsistency that exists in analyzing preference with multiple values as well as making decisions associated with typical cases, so that we can create rules that are more easily understood by users with the usage of hybrid models like this. Besides, rules based on the dominance relation are much more suitable than those based on the indiscernibility relation when they are used to classify new objects [20]. Because we may often meet preference information when handling with economic, managing, or financial decision-making problems, the hybrid of rough sets and dominance relation can enlarge the usage of rough set models in economic, management, or financial fields. Greco and other [4] scholars have put forward rough set model and its extended models based on dominance relation by replacing the indiscernibility relation with the dominance relation, and these models can do well with the possible inconsistency that exists in analyzing preference with multiple attributes as well as making decisions related to typical cases.

The granulation structures [14, 21] used by both rough set theory and neighbourhood systems, and the corresponding approximation structures, are studied.'

Life insurance is characterized by its long period ranging from 15 to 30 years in individual documents, and from one year to 45 years in the collective insurance. Premiums paid in the early years would be technically more than what should be paid despite the fact that the insured persons have paid equal premiums for the duration of insurance. These extra and accumulated sums are collected to be what is called the accounting reserve which is invested by the best means. The surplus or deficit amount in the formation of this reserve in life insurance is influenced by numerous technical elements related to compensation due for the deaths, rates of investment and productivity and administrative expenses of return and profits to be distributed to policyholders. It is possible after the cessation of the insured to pay premiums to cancel, modify, determine , reduce the amount or re- effect the recoverable amount of the document to secure a discount after the cancellation or waived , or even by borrowing , contrary to what is the case in the documents of property insurance and responsibility .

In most countries of the world, the law requires entitlement policyholder to obtain liquidation value often after paying a minimum premium of about three annual premiums. In the case of paying fewer or stopping to pay the premiums, the contract expires and becomes premiums paid go to the company issuing the insurance contract.

Studying Articles 760 and 762 of the Civil Law reveals that the Egyptian legislator did not give the insured any rights to the liquidation or reduction if the document if its three annual premiums are not paid. If the insured stops paying premiums during the first three years of life document for any reason, the document is cancelled after his warning and demanding repayment premiums due, and he does not have the right of returning to the company.

The problem of the research points out that the phenomenon of cancellations is one of the negative factors that limit the effectiveness of the insurance companies because it stops premiums flows which affects the potential contribution to invest and build up reserves. On the other hand, this phenomenon has a further negative impact on the holders of documents in general and the public insured in particular. Moreover, insurance companies should pay more attention to the phenomenon of cancellations and find appropriate solutions, particularly in developing countries, due to the increasing number abolition to the degree that it has become concrete for the number of the new production.

The contract is cancelled by the two parties. The insured can cancel the document on request. The insurer, however, can cancel only under certain conditions and must give the insured advanced written notice. In either case, the insured is entitled to a refund of the unearned premium on a pro-rated basis. The right of the insurer to cancel is limited by the provisions of the cancellation clause and may be done only under specified circumstances. Cancellation is the process of terminating coverage prior to the normal expiration date.

The following Table .1 shows the number of existing documents, insurance amounts, the number of canceled documents in addition to the cancellation of Misr life insurance in the Egyptian insurance market rate.

Table .1 Egypt company data for life insurance in the period (2007/ 2008 2014/2015)

Cancellation rate %	The number of canceled policies	Amounts of insurance	Number of applicable policies	the year
5.16	31914	12580567	618494	2008/2007
3.86	24452	14439741	633484	2009/2008
3.33	22661	16964965	668497	2010/2009
4.28	30333	19753633	708707	2011/2010
3.10	21970	21662418	708707	2012/2011
2.91	21182	23972591	727894	2013/2012
2.97	21618	27103785	727894	2014/2013
3.79	31864	30359745	840738	2015/2014

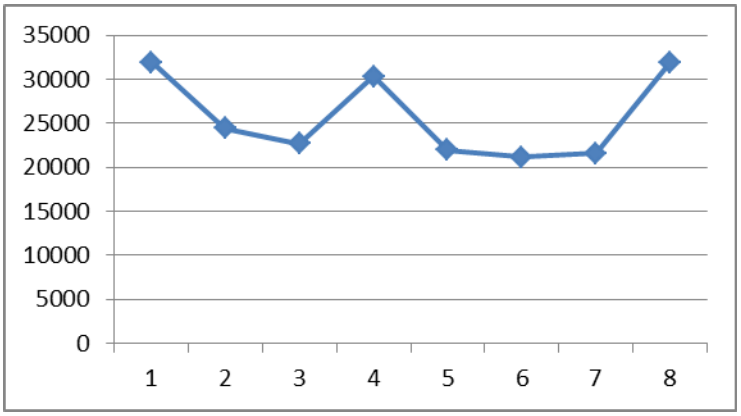

Fig. 1 shows the number of canceled documents in the period (2007/ 2008 2014/2015)

Given the above Fig. 1, we find that the number of canceled documents is rising up and oscillatory. Also, in Table (1), the rates of cancellations in constant fluctuation which requires searching the problem of cancellations of these documents , and identify the most important variables that affect the rates of cancellations and so as to reach solutions that lead to the reduction of the number of canceled documents and find solutions to them through the use of Rough model .

2. Preliminaries

2.1 Rough set theory [15] is a kind of mathematical method driven by practical needs to deal with vagueness and uncertainty. The main idea of the rough set theory is based on the indiscernibility relation that every object is associated with a certain amount of information and the object can only be expressed by means of some obtained information. Therefore, objects with the same or similar information can be indiscernible with respect to the available information. The indiscernible blocks formed by indiscernible objects of the universe are called the element sets or the knowledge granularity. Based on the knowledge used in approximation, the universe is divided into some element sets of the indiscernible objects that are based on conditional attribute sets, and the element sets are regarded as approximate "knowledge granularity"; the knowledge used in approximation would divide the universe into the decision-making class derived from the decision attribute sets, on this basis, and one kind of knowledge can be used to approximate the other. When approximation is inaccurate, roughness would appear.

The research also aims to use the Rough model to identify the determinants of cancellations in the life insurance of the Egyptian insurance market, so as to reach solutions that lead to the reduction of the number of canceled policies and find solutions to them.

This research derives its significance from being an attempt to figure out the size of cancellations in the life insurance policies issued by Misr Life Insurance.

2.2 Information Systems and Classification [8]

The starting point of the rough set is a data set, which is usually organized into a table, and it is called information systems or databases. The basic operators in rough set theory are lower and upper approximations, and they are used to define the total dependence and the part dependence of the property in the databases. Data reduct is a very important concept in rough set theory.

2.3 Information Systems and Indiscernibility Relation [10]

Knowledge representation is realized by information systems in the rough set model, and the form of information systems is the two-dimensional table showing the relationship between objects and values of an attribute, which is similar to the relational database. When classification is conducted according to the characteristics of the object, each attribute value is associated with one group of the value set, and when a value of each attribute is selected, a description of an object is given.

Definition 2. 3.1[11] Suppose that $S = (U, A, V, f)$ is an information system, and also can be called the knowledge representation system, where $U = \{U_1, U_2, \ldots, U_n\}$ is a finite nonempty set, and named universe object space; $A = \{a_1, a_2, \ldots, a_n\}$ is the finite nonempty set of attributes. $V = \cup V_a$, where $a \in A$, V_a is the domain of the attribute a; and f: $U \times A \rightarrow V$ is an information function, for $\forall a \in A$, $\forall x \in U$, and $f(x, a) \in V_a$, it appoints the attribute value of each object in U.

Definition 2. 3.2[18] Suppose $\forall a \in A$, $\forall x \in U$, and $f(x, a) \in V_a$; to every subset $\phi \neq P \subseteq A$, its indiscernibility relation I in U is defined as

$I = \{(x, y) \in U \times U : f(x,q) = f(y,q), \forall q \in P\}$, If $(x, y) \in I$, then x and y are indiscernible. Obviously, such definition of the indiscernibility relation is an equivalence relation, that is to say, the indiscernibility relation is characterized by reflexivity, symmetry, and transitivity. The equivalence class containing the object x is marked as $I(x)$. The equivalence class of indiscernibility relation is called as the element set or the atom in S, while the empty set is also the element of approximation space S. Equivalence class corresponds to the expressions of knowledge granularities, which is the basis of the main concepts in rough set, such as approximation, dependence, reduct, and so on.

2.4 Rough set and Approximations of Set[15]

The definition of set in the rough set theory is related with the available information (knowledge) and the understanding of the relevant universe elements. In other words, we can see universe elements with the available information of the universe elements. Therefore, two elements described by the same information are indiscernible.

The rough set method is to make a more identifiable data model based on the assumption of reducing the accuracy of data representation. The core assumption of "rough set" is that the knowledge is embodied in the ability of classification. That is to say, the rough set method is regarded as a standardized framework to find facts from incomplete data, whose result is expressed by the classification models or rules obtained from example sets.

Rough set theory embeds the knowledge to be classified into a set and makes it a part of the set. According to existing knowledge, there are three different situations as to whether an object $x \in U$ belongs to the set $X \subseteq U$, and they are as follows:

(1) Object x absolutely belongs to X;

(2) Object x absolutely does not belong to X;

(3) Object x may belong to X and may not belong to X.

On the basis of the above analysis, we put forward the two most important concepts in rough set theory: lower approximation and upper approximation.

Definition 2.4.1[17] Given a knowledge representation system $S = (U, A, V, f)$, $P \subseteq A$, $X \subseteq U$, $x \in U$, the lower approximation, upper approximation, negative region, and boundary of set X regarding I, respectively, are

$R_*(X) = \cup \{x \in U: I(x) \subseteq X\}$, $R^*(X) = \cup \{ x \in U: I(x) \cap X \neq \emptyset\}$, $neg(X) = \cup \{ x \in U: I(x) \cap X = \emptyset\}$, and $BN(X) = R^*(X) - R_*(X)$, will be referred to the B-boundary region of X. If the boundary region of X is the empty set, i.e., $BN(X) = \emptyset$, then the set X is crisp with respect to BNB(X); in the opposite case, i.e., if $BN(X) \neq \emptyset$, the set X is rough with respect to BN(X).

The lower approximation of set X is actually the largest union of all the objects that surely belong to X, and is also called the positive region of X denoted as $pos(X)$; the negative region of X denoted as $neg(X)$ is the set that is combined with all the objects that definitely do not belong to X, thus $neg(X) = U - R^*(X)$; the upper approximation denoted as $R^*(X)$ of X is consists of all nonempty sets of equivalence classes that intersect with X, that is to say, it is the smallest set that consists of the objects that may belong to X, apparently,

$U = R^*(X) \cup neg(X)$, while the boundary of X as, BN(X)is the difference between the upper approximation and lower approximation of the set X. The boundary region cannot be judged, for example, in terms of the attribute set P; there is no object in boundary region that can be classified into X or $-X$. If the BN(X) is an empty set, then the set X is an exact set with respect to I; otherwise, if the BN(X) is not an empty set, then the set X is rough set with respect to I.

3. Reduction and Core [8]

Every concept in the knowledge base can be only expressed in terms of basic categories. On the other hand, every basic category is built up of some elementary categories. The concepts of core and reduce are two fundamental concepts of the rough sets theory in the case of attributes and knowledge. The core is the common part of all reduces. The set of all indispensable attributes is called core. Logical rules derived from experimental data may be used to support new reduction. We often face a question whether we can remove some data from a data table preserving its basic properties, that is, whether a table contains some superfluous data. The proposed algorithm introduces indeterminacy by removing conditional attributes in a controlled manner. The selection of attributes to be removed is made from the factors in the discernibility function, thereby removing information needed to discern classes in the original information system.

4. Basic concepts

Definition 4.1 [2] If (U, A, V, f) is an information system defines an information function $f: U \rightarrow V$, where A is the set of attributes, V is the domain of the particular attributes in which the values V are real numbers. We define a relation R_{a_i} for each attribute, as follows: $x\ R_{a_i}\ y$ iff : $|a_i(x) - a_i(y)| \leq \varepsilon$, where ε is determined by an expert in the field. For example, if the information is from the medical field, the expert is a person interested in medicine and making in the problem. Thus for each $a_i \in A$, and O is a finite set, we can get a classification O/R_{a_i} which is $\{xR_{a_i} : x \in O\}$, $xR_{a_i} = \{y : |a_i(x) - a_i(y)| \leq \varepsilon$.

Definition 4.2 [16] Let R be a family of equivalence relation and let $A \in R$. Then we will say that A is dispensable in R if IND (R) = IND(R-{A}), (Reduce). But it was remarked IND (R) ≠ IND(R-{A}), then we will say that A is indispensable in R (Core).

5. Information Data

5.1. Data

In this article, we introduce a new reduction on the information technology and how areas of application of topology in the modern theory of rough set using information and data on an issue. We start the application with an appropriate information system by translating the values of the attributes $\{A_1, A_2, \ldots, A_8\}$ into qualitative terms. The data in question concern cancellation of insurance data, where 331 insurance data. The AAs are described in terms of eight attributes: A_1 = Amount insured, A_2 = premiums, A_3 = Age, A_4 = Occupation, A_5 = Address, A_6 = Method, A_7 = The period of insurance, and A_8 = The number of times the payment of the premium. The application can be described as follows; $S = \{x_1, x_2, \ldots, x_{331}\}$ denotes 8 listed attributes $\{A_1, A_2, \ldots, A_8\}$ into qualitative terms, as given in Table 2.

Table.2 Information system

Name of attribute Levels	Variable type	Levels	Codes related to their levels
Amount insured=A1	Continuous	Range 5000 to 30000	
Premiums=A2	Continuous	Range 20 to 359	
Age=A3	Continuous	Range 12 to 45	

Occupation=A4	Ordinal	Worker Student Academic Employee Housewife	1 2 3 4 5
Address=A5	Descriptive	Algharbiuh Albuhayra Kafr El Sheikh Alddaqahli alqilyubia	1 2 3 4 5
Method=A6	Ordinal	Annual midterm Third of annual Quarterly Monthly	1 2 3 4 5
The period of insurance= A7	Discrete	range 5 to 30	
A8=The number of times the payment of the premium	Discrete	range 1 to 16	

The coded information system is given by converting each attribute value to a value from 0 to 1 as follows: $V_{new} = (V_{old} - V_{min}) / (V_{max} – V_{min})$, where V_{max} and V_{min} are the maximum and minimum value for each attribute and by dividing the interval [0 , 1] into 5 parts as follows [0 , 0.20] =1, (0.20 , 0.40] =2 …, (0.80 , 1]= 5. The previous decision table contains attributes A_1, A_2, …, A_8 describes the belongingness of 331-objects.

5.2 Applying Rough sets for data

Based on the data presented, we can apply the stages of rough sets for acquiring knowledge about insurance data
(this knowledge presented in logical decision rules which declare the general description for the knowledge contained in the data), hence finding the minimal number of variables necessary for predicting the value of data.

After coding, if the eight attributes A_1, A_2, …, A_8 are taken into account, 313- elementary sets can be constructed. Table 3 gives the number of elementary sets after leaving out one of the attributes. For instance, if only A_1 is removed 268- elementary sets are distinguished, if A_2 is removed 293- elementary sets, if we remove attribute A_1, A_2, …, A_8, the number of elementary sets becomes smaller, whereas attributes A_1, A_2, …, A_8 are indispensable, as follows in Table 3.

Table 3. Removing Attributes

	Removing Attributes								
Number of elementary sets	None	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
	313	268	293	265	263	267	279	278	271

According A_1 is a decision, we get the result as shown in Table 4:

Table.4 According A_1 is a decision

Code	Class	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	5000-10000	134	107	151	107	44	117	71%
d=2	10000-15000	114	85	151	85	66	117	56%
d=3	15000-20000	29	19	46	19	27	222	41%
d=4	20000-25000	26	21	46	21	25	222	45%
d=5	25000-30000	10	8	15	8	7	253	53%
	Summation	313						

We found that all classes of the data as shown in table 4, when A_1 is a decision. The A_1-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough sets analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_1

For A_1 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A11)|$= 107, Upper approximation is $|U(A11)|$= 151, Positive region is $|Pos(A11)| = |L(A11)|$= 107, Boundary region is $|BN(A11)|$= $|U(A11)|$- $|L(A11)|$= 44, Negative region is $|NEG(A11)|$= $|S| - |U(A11)|$ = 268 - 151 = 117,

Accuracy of approximation μ $(A11)$ = $|L(A11)|$ / $|U(A11)|$ = 107/151 =71%.

For A_1 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A12)|$= 85, Upper approximation is $|U(A12)|$= 151, Positive region is $|Pos(A12)| = |L(A12)|$= 85, Boundary region is $|BN(A12)|$= $|U(A12)|$- $|L(A12)|$= 66,

Negative region is $|NEG(A12)|$= $|S| - |U(A12)|$ = 268 - 151 = 117

Accuracy of approximation μ $(A12)$ = $|L(A12)|$ / $|U(A12)|$ = 85/151 =56%.

For A_1 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A13)|$= 19, Upper approximation is $|U(A13)|$= 46, Positive region is $|Pos(A13)| = |L(A13)|$= 19, Boundary region is $|BN(A13)|$= $|U(A13)|$- $|L(A13)|$= 27,

Negative region is $|NEG(A13)|$= $|S| - |U(A13)|$ = 268 - 46 = 222,

Accuracy of approximation μ $(A13)$ = $|L(A13)|$ / $|U(A13)|$ = 19/46 =41%.

For A_1 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A14)|$= 21, Upper approximation is $|U(A14)|$= 46, Positive region is $|Pos(A14)| = |L(A14)|$= 21, Boundary region is $|BN(A14)|$= $|U(A14)|$- $|L(A14)|$= 25,

Negative region is $|NEG(A14)|$= $|S| - |U(A14)|$ = 268 - 46 = 222, Accuracy of approximation μ $(A14)$ = $|L(A14)|$ / $|U(A14)|$ = 21/46 =45%.

For A_1 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A15)|$= 8, Upper approximation is $|U(A15)|$= 15, Positive region is $|Pos(A15)|$ = $|L(A15)|$= 8,

Boundary region is $|BN(A15)|$= $|U(A15)|$- $|L(A15)|$= 7,

Negative region is $|NEG(A15)|$= $|S| - |U(A15)|$ = 268 - 15 = 253,

Accuracy of approximation μ $(A14)$ = $|L(A14)|$ / $|U(A14)|$ = 8/15 =53%.

According to Table 4, we found that most of the individuals who canceled their life insurance policies are the few amounts of insurance policies (their values ranged from 5000 to 10000 Egyptian pounds) for sample size equal 134 policy holders at 71% accuracy degree, while the life insurance policies that ranged from 10000 to 15000 come secondly with 58%accuracy degree for 114 policy holders as a research sample**.**

According A_2 is a decision; we get the result as shown in Table 5:

Table. 5 According A_2 is a decision

Code	Class	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	20-50	99	205	240	205	53	53	85%
d=2	50-100	152	48	73	48	25	220	65%
d=3	100-200	47	11	15	11	4	278	73%
d=4	200-250	5	9	11	9	2	282	81%
d=5	250-260	10	5	9	5	4	284	56%
	Summation	313						

We found that all classes of the data as shown in table 5, when A_2 is a decision. The A_2-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_2

For A_2 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A21)|$= 205, Upper approximation is $|U(A21)|$= 240, Positive region is $|Pos(A21)|$ = $|L(A21)|$= 205,

Boundary region is $|BN(A21)|$= $|U(A21)|$- $|L(A21)|$= 35,

Negative region is $|NEG(A21)|$= $|S| - |U(A21)|$ = 293 - 240 = 53,

Accuracy of approximation μ $(A21)$ = $|L(A21)|$ / $|U(A21)|$ = 205/240 =85%.

For A_2 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A22)|$= 48, Upper approximation is $|U(A22)|$= 73, Positive region is $|Pos(A22)|$ = $|L(A22)|$= 48,

Boundary region is $|BN(A22)|$= $|U(A22)|$- $|L(A22)|$= 25,

Negative region is $|NEG(A22)|$= $|S| - |U(A22)|$ = 293 - 73 = 220,

Accuracy of approximation μ $(A22)$ = $|L(A22)|$ / $|U(A22)|$ = 48/73 =65%.

For A_2 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A23)|$= 11, Upper approximation is $|U(A23)|$= 15, Positive region is $|Pos(A23)|$ = $|L(A23)|$= 11,

Boundary region is $|BN(A23)|$= $|U(A23)|$- $|L(A23)|$= 4,

Negative region is $|NEG(A23)|$= $|S| - |U(A23)|$ = 293 - 15= 278,

Accuracy of approximation μ $(A23)$ = $|L(A23)|$ / $|U(A23)|$ = 11/15 =73%.

For A_2 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A24)|$= 9, Upper approximation is $|U(A24)|$= 11, Positive region is $|Pos(A24)|$ = $|L(A24)|$= 9, Boundary region is $|BN(A24)| = |U(A24)| - |L(A24)| = 2$,

Negative region is $|NEG(A24)| = |S| - |U(A24)| = 293 - 11 = 282$,

Accuracy of approximation $\mu(A24) = |L(A24)| / |U(A24)| = 9/11 = 81\%$.

For A_2 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A25)| = 5$, Upper approximation is $|U(A25)| = 9$, Positive region is $|Pos(A25)| = |L(A25)| = 5$,

Boundary region is $|BN(A25)| = |U(A25)| - |L(A25)| = 4$,

Negative region is $|NEG(A25)| = |S| - |U(A25)| = 293 - 9 = 284$,

Accuracy of approximation $\mu(A25) = |L(A25)| / |U(A25)| = 5/9 = 56\%$.

Based on table 5, we found that most of the individuals who canceled their life insurance policies are the policies with small premiums (the premiums are ranged from 20 to 50 Egyptian pounds) for 99 policy holders as sample size at accuracy degree 85%, while the next category are the policyholders who canceled their life insurance policies (their premiums are ranged from 200 to 250 Egyptian pounds) with 81%accuracy degree for 5 policy holders as a research sample.

According A_3 is a decision, we get the result as shown in Table 6:

Table.6 Decision of attribute A_3

Code	Class	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	12-20	19	12	14	12	2	251	85%
d=2	20-25	56	63	114	63	51	151	55%
d=3	25-35	130	66	97	66	31	168	68%
d=4	35-40	87	76	127	76	51	138	60%
d=5	40-45	21	21	57	21	36	208	37%
	Summation	313						

We found that all classes of the data as shown in table 5, when A3 is a decision. The A_3-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_3

For A_3 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A31)| = 12$, Upper approximation is $|U(A31)| = 14$, Positive region is $|Pos(A31)| = |L(A31)| = 12$,

Boundary region is $|BN(A31)| = |U(A31)| - |L(A31)| = 2$,

Negative region is $|NEG(A31)| = |S| - |U(A31)| = 265 - 14 = 251$,

Accuracy of approximation $\mu(A31) = |L(A31)| / |U(A31)| = 12/14 = 85\%$.

For A_3 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A32)| = 63$, Upper approximation is $|U(A32)| = 114$, Positive region is $|Pos(A32)| = |L(A32)| = 63$,

Boundary region is $|BN(A32)| = |U(A32)| - |L(A32)| = 51$,

Negative region is $|NEG(A32)| = |S| - |U(A32)| = 265 - 114 = 151$,

Accuracy of approximation $\mu(A32) = |L(A32)| / |U(A32)| = 63/114 = 55\%$.

For A_3 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A33)| = 66$, Upper approximation is $|U(A33)| = 97$, Positive region is $|Pos(A33)| = |L(A33)| = 66$,

Boundary region is $|BN(A33)| = |U(A33)| - |L(A33)| = 31$,

Negative region is $|NEG(A33)| = |S| - |U(A33)| = 265 - 97 = 168$,

Accuracy of approximation $\mu(A33) = |L(A33)| / |U(A33)| = 66/97 = 68\%$.

For A_3 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A34)|$= 76, Upper approximation is $|U(A34)|$= 127, Positive region is $|Pos(A34)| = |L(A34)|$= 76,

Boundary region is $|BN(A34)|= |U(A34)|- |L(A34)|$= 51,

Negative region is $|NEG(A34)|= |S| - |U(A34)| =$ 265 - 127 = 138,

Accuracy of approximation $\mu\ (A34) = |L(A34)|\ /\ |U(A34)|$ = 76/127 =60%.

For A_3 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A35)|$= 21, Upper approximation is $|U(A35)|$= 57,

Positive region is $|Pos(A35)| = |L(A35)|$= 21,

Boundary region is $|BN(A35)|= |U(A35)|- |L(A35)|$= 36,

Negative region is $|NEG(A35)|= |S| - |U(A35)| =$ 265 - 57 = 208,

Accuracy of approximation $\mu\ (A35) = |L(A35)|\ /\ |U(A35)|$ = 21/57 =37%.

Table 6 showed that most of the individuals who canceled their life insurance policies are from the small ages (their ages are ranged from 12 to 20 years) with 85% accuracy degree, while the next category is youth from 25 to 35 years with accuracy degree 86%.

According A_4 is a decision, we get the result as shown in Table 7:

Table.7 Decision of attribute A_4

Code	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	19	12	34	12	22	229	35%
d=2	146	113	177	113	64	86	63%
d=3	35	27	48	27	21	215	56%
d=4	68	50	106	50	56	157	47%
d=5	45	33	67	33	34	196	49%
Summation	313						

We found that all classes of the data as shown in table 7, when A4 is a decision. The A_4-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_4

For A_4 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A41)|$= 12, Upper approximation is $|U(A41)|$= 34, Positive region is $|Pos(A41)| = |L(A41)|$= 12,

Boundary region is $|BN(A41)|= |U(A41)|- |L(A41)|$= 22,

Negative region is $|NEG(A41)|= |S| - |U(A41)| =$ 263 - 34 = 229,

Accuracy of approximation $\mu\ (A41) = |L(A41)|\ /\ |U(A41)|$ = 12/34 =35%.

For A_4 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A42)|$= 113, Upper approximation is $|U(A42)|$= 177, Positive region is $|Pos(A42)| = |L(A42)|$= 113,

Boundary region is $|BN(A42)|= |U(A42)|- |L(A32)|$= 64,

Negative region is $|NEG(A42)|= |S| - |U(A42)| =$ 263 - 177= 86,

Accuracy of approximation $\mu\ (A42) = |L(A42)|\ /\ |U(A42)|$ = 113/177 =63%.

For A_4 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A43)|$= 27, Upper approximation is $|U(A43)|$= 48, Positive region is $|Pos(A43)| = |L(A43)|$= 27,

Boundary region is $|BN(A43)|= |U(A43)|- |L(A43)|$= 21,

Negative region is $|NEG(A43)|= |S| - |U(A43)| =$ 263 - 48= 215,

Accuracy of approximation $\mu\ (A43) = |L(A43)|\ /\ |U(A43)|$ = 27/48 =56%.

For A_4 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A44)|$= 50, Upper approximation is $|U(A44)|$= 106, Positive region is $|Pos(A34)| = |L(A44)|$= 50, Boundary region is $|BN(A44)|= |U(A44)|- |L(A44)|$= 56, Negative region is $|NEG(A44)|= |S| - |U(A44)| =$ 263 - 106 = 157,

Accuracy of approximation $\mu\ (A44) = |L(A44)|\ /\ |U(A44)|$ = 50/106 =47%.

For A_4 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A45)|$= 33, Upper approximation is $|U(A45)|$= 67, Positive region is $|Pos(A45)| = |L(A45)|$= 33, Boundary region is $|BN(A45)|= |U(A45)|- |L(A45)|$= 34, Negative region is $|NEG(A45)|= |S| - |U(A45)| =$ 263 - 67 = 196,

Accuracy of approximation $\mu\ (A45) = |L(A45)|\ /\ |U(A45)|$ = 33/67 =49%.

Table 7, illustrated that students come first as the most category canceled their life insurance policies with 63% accuracy degree for 146 students as a research sample, while the next category is the academic staff category 56% as accuracy degree.

According A_5 is a decision, we get the result as shown in Table 8:

Table.8 Decision of attribute A_5

Code	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	115	86	146	86	60	121	59%
d=2	23	14	34	14	20	233	41%
d=3	35	27	49	27	22	218	55%
d=4	101	74	138	74	64	129	53%
d=5	39	34	53	34	19	214	64%
Summation	313						

We found that all classes of the data as shown in table 8, when A_5 is a decision. The A_5-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_5

For A_5 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A51)|$= 86, Upper approximation is $|U(A51)|$= 146, Positive region is $|Pos(A51)| = |L(A51)|$= 86, Boundary region is $|BN(A51)|= |U(A51)|- |L(A51)|$= 60, Negative region is $|NEG(A51)|= |S| - |U(A51)| =$ 267 - 146 = 121,

Accuracy of approximation $\mu\ (A51) = |L(A51)|\ /\ |U(A51)|$ = 86/146 =59%.

For A_5 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A52)|$= 14, Upper approximation is $|U(A52)|$= 34, Positive region is $|Pos(A52)| = |L(A52)|$= 14, Boundary region is $|BN(A52)|= |U(A52)|- |L(A52)|$= 20, Negative region is $|NEG(A52)|= |S| - |U(A52)| =$ 267- 34= 233,

Accuracy of approximation $\mu\ (A52) = |L(A52)|\ /\ |U(A52)|$ = 14/34 =41%.

For A_5 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A53)|$= 27, Upper approximation is $|U(A53)|$= 49, Positive region is $|Pos(A53)| = |L(A53)|$= 27, Boundary region is $|BN(A53)|= |U(A53)|- |L(A53)|$= 22, Negative region is $|NEG(A53)|= |S| - |U(A53)| =$ 267 - 49= 218,

Accuracy of approximation $\mu\ (A53) = |L(A53)|\ /\ |U(A53)|$ = 27/49 =55%.

For A_5 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A54)|$= 74, Upper approximation is $|U(A54)|$= 138, Positive region is $|Pos(A54)| = |L(A54)|$= 74, Boundary region is $|BN(A54)| = |U(A54)| - |L(A54)|$= 64, Negative region is $|NEG(A54)| = |S| - |U(A54)|$ = 267 - 138 = 129,

Accuracy of approximation μ $(A54)$ = $|L(A54)|$ / $|U(A54)|$ = 74/138 =53%.

For A_5 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A55)|$= 34, Upper approximation is $|U(A55)|$= 53, Positive region is $|Pos(A55)| = |L(A55)|$= 34, Boundary region is $|BN(A55)| = |U(A55)| - |L(A55)|$= 19, Negative region is $|NEG(A55)| = |S| - |U(A55)|$ = 267 - 53 = 214,

Accuracy of approximation μ $(A55)$ = $|L(A55)|$ / $|U(A55)|$ = 34/55 =64%.

Table 8, demonstrated that Alqilyubia is the greatest governorates in life insurance policies cancelation with 64% accuracy degree, whereas Algharbiuh governorate comes next with 59% accuracy degree.

According A_6 is a decision, we get the result as shown in Table 9:

Table.9 Decision of attribute A_6

Code	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	8	5	13	5	8	266	38%
d=2	13	9	23	9	14	256	39%
d=3	2	1	3	1	2	276	33%
d=4	76	64	99	64	35	180	65%
d=5	214	186	239	186	53	40	78%
Summation	313						

We found that all classes of the data as shown in table 9, when A_6 is a decision. The A_6-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_6

For A_6 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A61)|$= 5, Upper approximation is $|U(A61)|$= 13, Positive region is $|Pos(A61)| = |L(A61)|$= 5, Boundary region is $|BN(A61)| = |U(A61)| - |L(A61)|$= 8, Negative region is $|NEG(A61)| = |S| - |U(A61)|$ = 279 - 13 = 266,

Accuracy of approximation $\mu(A61)$ = $|L(A61)|$ / $|U(A61)|$ = 5/13 =38%.

For A_6 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A62)|$= 9, Upper approximation is $|U(A62)|$= 23, Positive region is $|Pos(A62)| = |L(A62)|$= 9, Boundary region is $|BN(A62)| = |U(A62)| - |L(A62)|$= 14, Negative region is $|NEG(A62)| = |S| - |U(A62)|$ = 279 - 23= 256,

Accuracy of approximation μ $(A62)$ = $|L(A62)|$ / $|U(A62)|$ = 9/23 =39%.

For A_6 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A63)|$= 1, Upper approximation is $|U(A63)|$= 3, Positive region is $|Pos(A63)| = |L(A63)|$= 1, Boundary region is $|BN(A63)| = |U(A63)| - |L(A63)|$= 2, Negative region is $|NEG(A63)| = |S| - |U(A63)|$ = 279 - 3= 276,

Accuracy of approximation μ $(A63)$ = $|L(A63)|$ / $|U(A63)|$ = 1/3 =33%.

For A_6 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A64)|$= 64, Upper approximation is $|U(A64)|$= 99, Positive region is $|Pos(A64)| = |L(A64)|$= 64,

Boundary region is $|BN(A64)|= |U(A64)|- |L(A64)|$= 35,

Negative region is $|NEG(A64)|= |S| - |U(A64)| =$ 279 - 99 = 180,

Accuracy of approximation $\mu(A64) = |L(A64)| / |U(A64)|$ = 64/99 =65%.

For A_6 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A65)|$= 186, Upper approximation is $|U(A65)|$= 239, Positive region is $|Pos(A65)| = |L(A65)|$= 186,

Boundary region is $|BN(A65)|= |U(A65)|- |L(A65)|$= 53,

Negative region is $|NEG(A65)|= |S| - |U(A65)| =$ 279 - 239 = 40,

Accuracy of approximation $\mu(A65) = |L(A65)| / |U(A65)|$ = 186/239 =78%.

Our results approved that the majority of policyholders who cancel their life insurance pay their premiums monthly with 78% accuracy degree for 214 as a sample for this research, while the policyholders who pay their premiums quarterly come next with 65% accuracy degree for sample equal 76 policy holders as shown in Table 9.

According A_7 is a decision, we get the result as shown in Table 10:

Table.10 Decision of attribute A_7

Code	Class	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	5-10	6	6	8	6	2	270	75%
d=2	10-15	60	57	69	57	12	209	82%
d=3	15-20	72	61	98	61	37	180	62%
d=4	20-25	160	131	177	131	46	101	74%
d=5	25-30	15	13	24	13	11	254	54%
	Summation	313						

We found that all classes of the data as shown in table 10, when A_7 is a decision. The A_7-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_7

For A_7 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A71)|$= 6, Upper approximation is $|U(A71)|$= 8, Positive region is $|Pos(A71)| = |L(A71)|$= 6,

Boundary region is $|BN(A71)|= |U(A71)|- |L(A71)|$= 2,

Negative region is $|NEG(A71)|= |S| - |U(A71)| =$ 278 - 8 = 270,

Accuracy of approximation $\mu(A71) = |L(A71)| / |U(A71)|$ = 6/8 =75%.

For A_7 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A72)|$= 57, Upper approximation is $|U(A72)|$= 69, Positive region is $|Pos(A72)| = |L(A72)|$= 57,

Boundary region is $|BN(A72)|= |U(A72)|- |L(A72)|$= 12,

Negative region is $|NEG(A72)|= |S| - |U(A72)| =$ 278 - 69= 209,

Accuracy of approximation $\mu(A72) = |L(A72)| / |U(A72)|$ = 57/69 =82%.

For A_7 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A73)|$= 61, Upper approximation is $|U(A73)|$= 98, Positive region is $|Pos(A73)| = |L(A73)|$= 61,

Boundary region is $|BN(A73)|= |U(A73)|- |L(A73)|$= 37,

Negative region is $|NEG(A73)|= |S| - |U(A73)| =$ 278 - 98= 180,

Accuracy of approximation $\mu(A73) = |L(A73)| / |U(A73)|$ = 61/98 =62%.

For A_7 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A74)|$= 131, Upper approximation is $|U(A74)|$= 177, Positive region is $|Pos(A74)| = |L(A74)|$= 131, Boundary region is $|BN(A74)|= |U(A74)|- |L(A74)|$= 46, Negative region is $|NEG(A74)|= |S| - |U(A74)| =$ 278 - 177 = 101,

Accuracy of approximation $\mu\ (A74) = |L(A74)|\ /\ |U(A74)|$ = 131/177 =74%.

For A_7 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A75)|$= 13, Upper approximation is $|U(A75)|$= 24, Positive region is $|Pos(A75)| = |L(A75)|$= 13, Boundary region is $|BN(A75)|= |U(A75)|- |L(A75)|$= 11, Negative region is $|NEG(A75)|= |S| - |U(A75)| =$ 278 - 24 = 254,

Accuracy of approximation $\mu(A75) = |L(A75)|\ /\ |U(A75)|$ = 13/24 =54%.

According A_8 is a decision, we get the result as shown in Table 11:

Our findings revealed that the highest category of individuals who canceled their life insurance policies are those hold insurance policies with 5 to 10 years old with 75% accuracy degree for sample equal 6 policy holders, while the policies with duration ranged from 20 to 25 years come secondly in arrangement with74% accuracy degree for sample equal 160. Moreover, the policies with duration ranged from 10 to 15years come thirdly in the arrangement with74% accuracy degree for sample equal 60 as shown in Table 10.

According A_8 is a decision, we get the result as shown in Table 11:

Table.11 Decision of attribute A_8

Code	Class	Frequency	Lower approximations	Upper approximations	Positive regions	Boundary regions	Negative regions	Accuracy of approximation
d=1	1-3	178	156	214	156	58	57	73%
d=2	4-6	60	40	82	40	42	189	49%
d=3	7-9	24	18	33	18	15	238	54%
d=4	10-12	29	19	40	19	21	231	47%
d=5	13-16	22	14	35	14	21	236	40%
	Summation	313						

We found that all classes of the data as shown in table 11, when $\boldsymbol{A_8}$ is a decision. The $\boldsymbol{A_8}$-decision is divided 5-classes as follows, d1=1, d2=2, d3=3, d4=4 and d5=5.

The next step of the rough set analysis is to find Lower approximation, Upper approximation, Positive, Boundary and Negative regions for the partition induced by indiscernibility relation IND(D)

Discussion- A_8

For A_8 (Set of objects which have the same decision (d1 = 1))

Lower approximation is $|L(A81)|$= 156, Upper approximation is $|U(A81)|$= 214, Positive region is $|Pos(A81)| = |L(A81)|$= 156, Boundary region is $|BN(A81)|= |U(A81)|- |L(A81)|$= 58, Negative region is $|NEG(A81)|= |S| - |U(A81)| =$ 271 - 214= 57,

Accuracy of approximation $\mu\ (A81) = |L(A81)|\ /\ |U(A81)|$ = 156/214 =73%.

For A_8 (Set of objects which have the same decision (d = 2))

Lower approximation is $|L(A82)|$= 40, Upper approximation is $|U(A82)|$= 82, Positive region is $|Pos(A82)| = |L(A82)|$= 40, Boundary region is $|BN(A82)|= |U(A82)|- |L(A82)|$= 42, Negative region is $|NEG(A82)|= |S| - |U(A82)| =$ 271 - 82= 189,

Accuracy of approximation $\mu\ (A82) = |L(A82)|\ /\ |U(A82)|$ = 40/82 =49%.

For A_8 (Set of objects which have the same decision (d = 3))

Lower approximation is $|L(A83)|$= 18, Upper approximation is $|U(A83)|$= 33, Positive region is $|Pos(A83)| = |L(A83)|$= 18, Boundary region is $|BN(A83)|= |U(A83)|- |L(A83)|$= 15, Negative region is $|NEG(A83)| = |S| - |U(A83)| = 271 - 33 = 238$,

Accuracy of approximation $\mu\ (A83) = |L(A83)| / |U(A83)| = 18/33 = 54\%$.

For A_8 (Set of objects which have the same decision (d = 4))

Lower approximation is $|L(A84)| = 19$, Upper approximation is $|U(A84)| = 40$, Positive region is $|Pos(A84)| = |L(A84)| = 19$,

Boundary region is $|BN(A84)| = |U(A84)| - |L(A84)| = 21$,

Negative region is $|NEG(A84)| = |S| - |U(A84)| = 271 - 40 = 231$,

Accuracy of approximation $\mu\ (A84) = |L(A84)| / |U(A84)| = 19/40 = 47\%$.

For A_8 (Set of objects which have the same decision (d = 5))

Lower approximation is $|L(A85)| = 14$, Upper approximation is $|U(A85)| = 35$, Positive region is $|Pos(A85)| = |L(A85)| = 14$,

Boundary region is $|BN(A85)| = |U(A85)| - |L(A85)| = 21$,

Negative region is $|NEG(A85)| = |S| - |U(A85)| = 271 - 35 = 236$,

Accuracy of approximation $\mu\ (A85) = |L(A85)| / |U(A85)| = 14/35 = 40\%$.

In Table 11, our results showed that policyholders who pay three premiums or less represent the highest class of life insurance policies cancellation with 73% accurate degree for 178 as a sample size, wherase the policyholders who paid from seven to nine premiums come next with 54% accurate degree for 24 life insurance policyholders.

Final Discussion

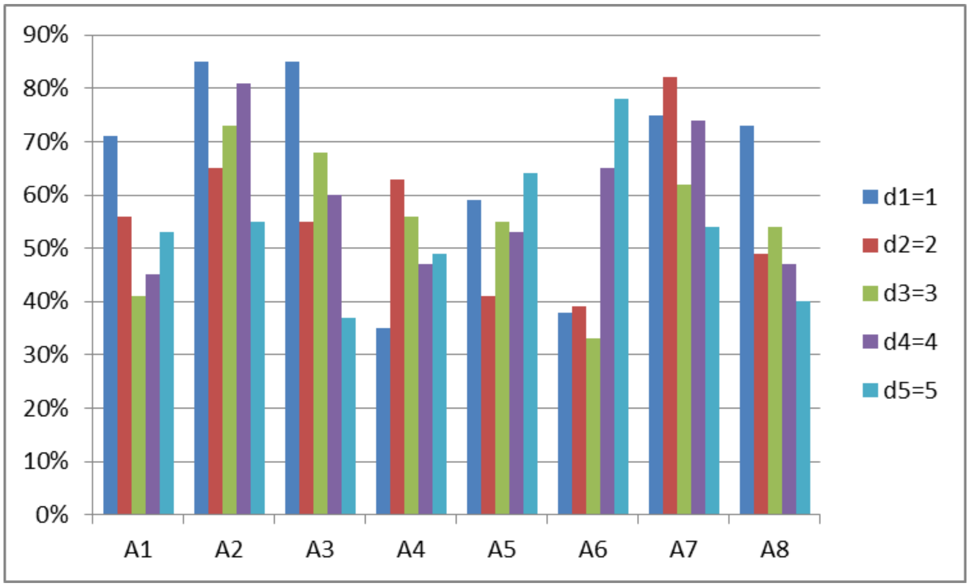

Fig (2) Final decision

It is clear from figure 2 that the attributes A_1, A_2, A_3, and A8, and when the decision is equal to one (d1=1) is given the value of super these attributes, while attributes A_4, A_7 and when the decision is equal to two (d2=2), we find that it gives great value to these attributes, As well as the attributes A_5, A_6 and when the decision is equal to five (d5=5), it result in great value for these attributes.

We studied the great values of the attributes because they are expressing people intention to cancel their insurance, and therefore we were able to determine which attributes are more associated with the cancellation behaviour.

6. Conclusion

Based on the analysis of the above tables, the accuracies of the information system determined the best attributes to be used from insurance data. This work is a new way to use the rough set where we got the best attributes from the insurance data as it is reflected in the final table (12) as follows.

Table. 12 Final decisions

	d1=1	d2=2	d3=3	d4=4	d5=5	Average decision
$\mu 1$	71%	56%	41%	45%	53%	52%
$\mu 2$	85%	65%	73%	81%	55%	71%
$\mu 3$	85%	55%	68%	60%	37%	59%
$\mu 4$	35%	63%	56%	47%	49%	49%
$\mu 5$	59%	41%	55%	53%	64%	54%
$\mu 6$	38%	39%	33%	65%	78%	48%
$\mu 7$	75%	82%	62%	74%	54%	69%
$\mu 8$	73%	49%	54%	47%	40%	52%

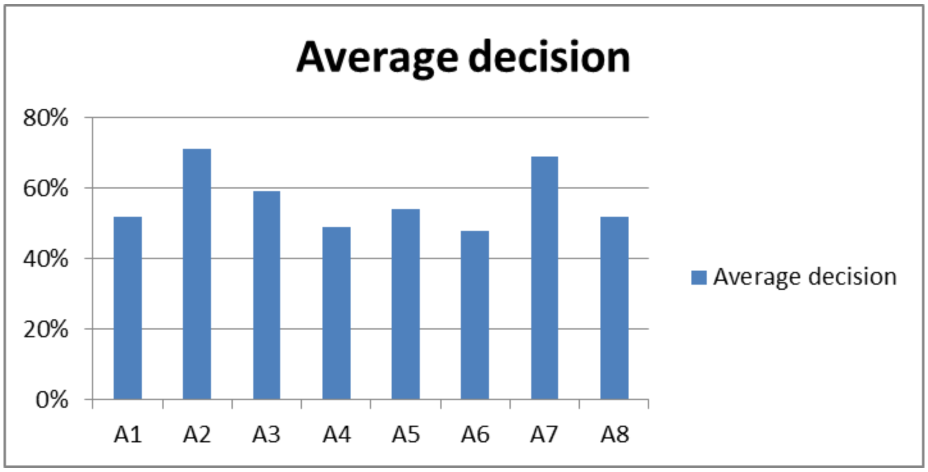

Fig.3 Average decision

It is obvious from Fig (3) that, more attributes to cancel insurance are A_2, A_7, A_3, respectively, where A_2 refers to the insurance premium value, A_7 refers to the period of insurance, A_3 indicate age. These results are realistic, so the expert in the field of insurances handles these things in order to get the best result.

The results of the rough set approach are presented in the form of classification or decision rules derived from a set of the previous application. This study provides a new insight into the problem of attribute reduction.

References

[1] Abd El-Monsef M M E, and Kilany N M. (2007), Decision analysis via granulation based on general binary relation, International Journal of Mathematics and Mathematical Sciences, Vol. 2007: 1-13. http://dx.doi.org/10.1155/2007/12714

[2] Alharthi T N, and Elsafty M A. (2016), Attribute topologies based similarity, Cogent mathematics journal, Vol. 3: 1-11. http://dx.doi.org/10.1080/23311835.2016.1242291

[3] Chan C C. (1998), A rough set approach to attribute generalization in data mining. Journal of Information Sciences, Vol. 107: 169–176. http://dx.doi.org/10.1016/S0020-0255(97)10047-0

[4] Creco S, Inuiguchi M, and Slowinski R. (2006), Fuzzy rough sets and multiple-premise gradual decision rules, International Journal of Approximate Reasoning, Vol. 41: 179–211. http://dx.doi.org/10.1016/j.ijar.2005.06.014

[5] Flapan E. (2000), When Topology meets Chemistry. Cambridge: Cambridge University press. http://dx.doi.org/10.1017/CBO9780511626272 K

[6] Greco S, Matarazzo B, and Slowinski R. (2001), Rough sets theory for multi-criteria decision analysis, European Journal of Operational

Research, Vol. 129: 1–47. http://dx.doi.org/10.1016/S0377-2217(00)00167-3

[7] Jhieh-Yu Shyng , Fang-Kuo Wang , Gwo-Hshiung Tzeng , (2007), Kun-Shan Wu,(2007), Rough Set Theory in analyzing the attributes of combination values for the insurance market, science direct journal, Expert Systems with Applications, Vol. 32: 56–64. http://dx.doi.org/10.1016/j.eswa.2005.11.002

[8] Kozae A M, Elsafty M, and Swealam M. (2012), Neighbourhood and Reduction of Knowledge, Advances in information Sciences and Service Sciences(AISS). Vol. 4: 247-253.http://dx.doi: 10.4156/AISS.vol4.issue1.32

[9] Kun-Shan Wu1, Fang-Kuo Wang and Jhieh-Yu Shyng ,(2013), Applied data mining techniques in insurance company: A comparative study of rough sets and decision tree, African Journal of Business Management, Vol. 7: 2309-2318. http://dx.doi.org/10.1109/ICSMC.2006.384912

[10] Lashin E F, and Medhat T. (2005), Topological reduction of information systems, Chaos, Solitons and Fractals, Vol. 25: 277–286. http://dx.doi.org/10.1016/j.chaos.2004.09.107

[11] Lashin E F, Kozae A M, Abo Khadra A A, and Medhat T. (2005), Rough set theory for topological spaces, International Journal of Approximate Reasoning, Vol. 40: 35-43. http://dx.doi.org/10.1016/j.ijar.2004.11.007

[12] Li H, Zhou X and Huang B. (2007), Attribute reduction in incomplete information systems based on connection degree rough set, *Computer Science (Ji Suan JiKe Xue)*, Vol. 34: 39-42. http://dx.doi.org/10.1109/CHICC.2006.4347103

[13] Li K, and Liu Y S. (2002), Rough set based attribute reduction approach in data mining, In: *Proceedings of 2002 International Conference on Machine Learning and Cybernetics*. Piscataway, NJ: IEEE Press, 60–63. http://dx.doi.org/10.1109/ICMLC.2002.1176709

[14] Pal S K, Shankar B U, and Mitra P. (2005), Granular computing, rough entropy and object extraction, *Pattern Recognition Letters*, Vol. 26: 2509–2517. http://dx.doi.org/10.1016/j.patrec.2005.05.007

[15] Pawlak Z. (1982), Rough sets, International journal of computer and information science, Vol. 11: 341-356. http://dx.doi.org/10.1007/BF01001956

[16] Pawlak Z. (2002), Rough sets and intelligent data analysis, Information Science, Vol. 147: 1–12. http://dx.doi.org/10.1016/S0020-0255(02)00197-4

[17] Pawlak Z, and Skowron A. (2007), Rough sets: some extensions, Information Science, Vol. 177: 41-73. http://dx.doi.org/10.1016/j.ins.2006.06.006

[18] Salama S. (2011), Some topological properties of rough sets with tools for data mining', International Journal of Computer Science, Vol. 8: 588-595. http://dx.doi.org/10.1007/978-3-642-21881-1_7

[19] Wang F H. (2005), On acquiring classification knowledge form noisy data based on the rough set, Expert Systems Application, Vol. 29: 49–64. http://dx.doi.org/10.1016/j.eswa.2005.01.005

[20] Yao Y Y. (1998), Relational interpretations of neighbourhood operators and rough set approximation operators, Information Sciences, Vol. 111: 239–259. http://dx.doi.org/10.1016/S0020-0255(98)10006-3

[21] Yao Y Y. (1999), Rough Sets, Neighborhood Systems, and Granular Computing, Proceedings of the 1999 lEEE Canadian Conference on Electrical and Computer Engineering Shaw Conference Center, Edmonton, Alberta, Canada. http://dx.doi.org/10.1109/CCECE.1999.804943

[22] Yao YY. (2007), Decision-theoretic rough set models, Proceedings of RSKT'07, LNAI, Vol. 4481: 1-12. http://dx.doi.org/10.1007/978-3-540-72458-2_1