CONSTRUCTION OF CLASSIFIERS IN GENERALIZED COVERING APPROXIMATION SPACES

Dariusz Małyszko

Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland

Abstract: Emerging intelligent information systems are pushing existing mathematical foundations into new directions. Generalized covering approximation spaces present abstract data model useful in development of new data analysis methods. The paper introduces construction of rough classifiers in generalized covering approximation spaces. The main idea comes from generation of rough coverings in feature space and calculation of rough covering descriptor. Data are divided into data blocks and each data block statistic and bounding block is calculated . Feature space is divided into feature blocks. For each data bounding block, its inclusion into feature block is calculated and rough covering descriptor is created. Rough covering descriptor is embedded in the generalized covering approximation spaces with standard, fuzzy and probabilistic coverings giving robust theoretical framework in design, implementation and application of classification algorithms.

Keywords: Generalized approximation spaces, covering approximation spaces, rough covering approximation spaces

1. Introduction

Modern computer systems are shifting into transformation to intelligent information systems requiring extension of existing mathematical theories into new directions. Rough set theory presents framework designed for handling imprecise, vague, incomplete, uncertain information. A fundamental methodology of the theory of rough sets has been grounded by means of division of universe object into classes defined by an equivalence relation. In order to extend this point of view, many research has been focused on presenting more generalized data models. Classical rough set theory based equivalence relations have been replaced by more general binary relations, coverings, fuzzy sets, neighborhood systems and general approximation framework.

Recently, the concept of neighborhood has been introduced to define and study different types of neighborhood-based covering rough sets. The covering-based rough

Advances in Computer Science Research, vol. 13, pp. 43-58, 2016.

sets is one of the most important extensions of the classical Pawlak rough sets. In the covering-based rough set theory, various kinds of rough sets were already defined and studied in the literature. The relation among different covering approximation operators has been presented in [1], [2]. In [3] a framework for the study of covering approximation operators by the element based, the granule based and the subsystem based definitions. Approximation operators in covering based rough sets are presented in [4]. Topological properties of covering rough sets are investigated in [5]. Fuzzy coverings rough sets introduced in [6] are linking covering rough set theory and fuzzy rough set theory.

The generalized rough sets in neighborhood system is another important extension of the classical Pawlak rough sets. Rough set extensions have been presented in [7]. Rough feature covering model creates coverings of universe in the generalized approximation spaces. Coverings create neighborhoods of objects, inclusion function is based upon inclusion of coverings.

Rough Extended Framework presented in [8], [9], [10] extensively developed method of data analysis based upon data structure inferred from metric relations in rough, fuzzy and probabilistic approach. The theory of rough sets and fuzzy sets have applied in many image analysis algorithms as described in [11] and can also be combined with various computational intelligence techniques.

Rough covering model embeds into generalized covering approximation spaces with coverings creating neighborhood system, inclusion function describes degree of inclusion for coverings. Established generalized approximation space model gives good foundations for data analysis. Feature coverings are based upon clustering and thresholding of feature space. Rough covering model incorporates standard, fuzzy and probabilistic data model, giving interoperability. The main contribution of the paper is presentation of precise framework for creation rough covering descriptor and algorithm for data classification that is embedded in generalized covering approximation spaces.

This paper has been structured in the following way. In Section 2. the introductory information about rough sets, approximation spaces and generalized covering approximation spaces has been presented. In Section 3. the covering types of generalized covering approximation spaces have been described. Construction of classifiers based upon rough descriptors in generalized approximation spaces is given in Section 4.. Next, concluding remarks and future research are given.

2. Generalized covering approximation spaces

In this Section, definition of generalized deep feature covering approximation space is given with introduction of inclusion function for this space. Next approximation neighborhood systems in these spaces are introduced that form rough approximations of feature coverings.

Definition 1 *Universe consists of objects, further referred to as universe objects. Universe object consists of vector of n-dimensional points by default in R^n space.*

Universe object is the vector of m points, $p = (p_1, \dots, p_m)$. Each point $p_i = (x_1, \dots, x_n)$ is described by n attributes or features from attribute set $A = \{a_1, \dots, a_n\}$. In case of the number of points of the universe object equal 1, the classical definition of the universe is obtained. Each attribute a_i has domain of possible values, forming real interval defined as $da_i = (min_{a_i}, max_{a_i})$. In that way, the concept of attribute space is introduced

Definition 2 *Attribute space or feature space $V = (da_i \times \cdots \times da_n)$ is the space of all possible feature values for each attribute, it means $da_i = (min_{a_i}, max_{a_i})$. Given the universe U, attribute space for this universe will be denoted as V(U).*

Corollary 1. *Both elements of the universe U and attribute space V(U) consist of n-dimensional points so they have the same domain and may be compared.*

Universe objects as the sets of n-dimensional points are divided into data blocks forming data covering. The same applied to attribute space V(U) as the sets of feature vectors divided into feature blocks forming feature coverings.

Definition 3 *A data block N = $\{p_1, \dots, p_b\}$ is a set of n-dimensional points of universe object. The order of a data block is the number of points of data block.*

Definition 4 *A feature block is a set of n-dimensional points. The order of a feature block is the number of points of feature block including infinite case. Feature blocks of universe objects are denoted as f=$\{p_1, \dots, p_b\}$ and feature blocks of attribute space are denoted as F=$\{p_1, \dots, p_b\}$.*

Definition 5 *Selection of subsets of universe creates universe covering and selection of subsets of universe object creates universe object covering.*

Theorem 1. *Each covering (subsets) of universe object of n-dimensional points defines feature blocks f.*

As universe objects consists of n-dimensional points, selecting some number of n-dimensional points fulfills the definition of feature block.

Theorem 2. *Each covering (subsets) of feature space V(U) defines feature blocks F.*

As feature space consists of n-dimensional points, selecting some number of n-dimensional points fulfills the definition of feature block.

Definition 6 *A feature block $F = \{p_1, \ldots, p_m\}$ has min, max and avg values determined as*

$$a(F) = avg(p_1, \ldots, p_m)$$
$$m(F) = min(p_1, \ldots, p_m)$$
$$M(F) = max(p_1, \ldots, p_m)$$

where values avg, min, max are calculated for each attribute over all n-dimensional points forming feature block.

Definition 7 *A hyperbox defines region in n-dimensional space determined by means two n-dimensional points, formally defined as the Cartesian product of intervals In an n-dimensional space of real numbers R^n, let $a, b \in R^n$ be given points such that $a < b$, i.e. $\forall_{1 \le i \le n}\ a_i < b_i$. A hyperbox H=(a,b) is the set of points satisfying condition*

$$H \subset R^n, x = \{x_1, \ldots, x_n\}, x \in H \iff \forall_{1 \le i \le n} \quad a_i < x_i < b_i$$

where x_j is j-th feature of x.

Theorem 3. *Each feature block F defines H hyperbox in R^n by assuming*

$$H = (m(F), M(F))$$

A direct consequence of theorem 1 is the fact

Corollary 2. *Each data covering defines H hyperboxes in R^n.*

A consequence of theorem 2 is the fact

Corollary 3. *Each feature covering defines H hyperboxes in R^n.*

Corollary 4. *Feature blocks F of attribute space form some kind of universe objects with accordance to definition 1.*

From corollary 4 results that feature blocks F can be compared to universe objects.

Definition 8 *Hyperbox $H = (a,b)$ has min, max and avg values determined as*

$$a(H) = (a+b)/2$$
$$m(H) = a$$
$$M(H) = b$$

In that way, that comparison between data blocks and feature blocks is possible on the base of inclusion of their respective hyperboxes. Selection of similar indiscernible objects creates elementary sets that describe our knowledge about the universe. Elementary sets are called precise sets. All other sets are considered to be rough, imprecise, vague. An information system IS = (U,A) consists of objects described by attributes. Attribute a_i is defined as a function $a_i : U \to V_{a_i}$ where V_{a_i} is a set of attribute values. The same applies to extended definition of universe objects from definition 1. Information system with objects divided into classes by reflexive, symmetric and transitive equivalence R on U is called an approximation space. Lower and upper approximations of any subset X of U are defined as two sets completely contained or partially contained in the equivalence classes. This approach have been generalized by extending equivalence relations to tolerance relations, similarity relations, binary relations leading into formulation of the concept of Generalized approximation spaces.

Definition 9 *A generalized approximation space is a tuple GAS = (U, N, υ) with N is a neighborhood function defined on U with values in the powerset $P(U)$ of U. The overlap function υ is defined on the Cartesian product $P(U) \times P(U)$ with values in the interval [0, 1] measuring the degree of overlap of sets.*

The lower GAS_* and upper GAS^* approximation operations can be defined in a GAS by

$$GAS_*(X) = \{x \in U : \upsilon(N(x), X) = 1\},$$
$$GAS^*(X) = \{x \in U : \upsilon(N(x), X) > 0\}.$$

Generalized approximation spaces present environments with specialized rough set models applied such as similarity based rough set model with reflexive neighborhood function and variable precision model with different thresholds definition in overlap functions.

Definition 10 *Let C be a family of nonempty subsets of U creating a covering of U. The ordered pair CAS = ($\mathcal{C}$, U) is called a covering approximation space.*

Definition 11 *Let every universe object x of U has a family of nonempty subsets of x creating a covering of x. The ordered pair DAS = ($\mathcal{D}$, U) is called a deep covering approximation space.*

Starting from covering approximation spaces and defining more specialized neighborhoods with overlap function we obtain the following definitions.

Neighborhood system N_S with two distinct coverings L, U for each universe object is called approximation neighborhood system.

Definition 12 *Let C be a family of nonempty subsets of attribute feature space V(U) associated with U creating a covering of U. The ordered pair CAS = ($\mathcal{C}$,V(U)) is called a feature covering approximation space.*

Definition 13 *Given covering approximation space CAS, approximation neighborhood system N_S assigns for each object x two coverings L, U containing that object $N_S = \{L, U \in \mathcal{C} : x \in L, U\}$*

Approximation neighborhood system R_S with two distinct coverings L, U for each universe object satisfying condition $L \subset U$ is called rough approximation neighborhood system R_S.

Definition 14 *In covering approximation space CAS, rough approximation neighborhood system $R_S(x) = \{L, U \in \mathcal{C} : x \in L, U\}$ defines two coverings L, U containing x, satisfying $L \subset U$.*

Definition 15 ***A generalized covering approximation space*** *is a system GCS = (U, $\mathcal{C}$, R_S, υ) where R_S is a rough approximation neighborhood system defined on U with covering $\mathcal{C}$, and υ is the overlap function measuring inclusion of coverings from $\mathcal{C}$.*

In order to apply coverings to universe objects the following notion of generalized deep covering approximation space has been introduced.

Definition 16 ***A generalized deep covering approximation space*** *is a system GDS = (U, $\mathcal{D}$, R_S, υ) where R_S is a rough approximation neighborhood system defined on universe objects of U with covering $\mathcal{D}$, and υ is the overlap function measuring inclusion of coverings from $\mathcal{D}$.*

In order to embed into covering approximation spaces additional coverings of feature space the notion of generalized feature covering approximation space has been given.

Definition 17 ***A generalized feature covering approximation space*** *is a system GFS = (U, $\mathcal{C}$, $\mathcal{F}$, R_S, υ) where R_S is a rough approximation neighborhood system defined on U with covering $\mathcal{C}$, covering of feature space $\mathcal{F}$ and υ is the overlap function measuring inclusion of coverings from $\mathcal{C}$ and $\mathcal{F}$.*

Generalized feature approximation spaces have coverings of universe and coverings of feature space. When coverings of universe are replaced by coverings of universe objects the following definition of generalized deep feature covering approximation space is obtained.

Definition 18 ***A generalized deep feature covering approximation space*** *is a system GFDS = (U, $\mathcal{D}$, $\mathcal{F}$, R_S, υ) where R_S is a rough approximation neighborhood system defined on universe objects U with coverings $\mathcal{D}$ of universe objects, covering $\mathcal{F}$ of feature space and υ is the overlap function measuring inclusion of two coverings from $\mathcal{D}$ and $\mathcal{F}$.*

Definition 19 ***Inclusion function*** *in generalized covering approximation spaces has been defined in the way presented below*

The degree of feature hyperbox inclusion represented by feature overlap function is defined as

$$v(H_i, H_j) = \begin{cases} 1 & \text{if } a(H_i) \in H_j \\ 0.5 & \text{if } H_i \cap H_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

with $a(H_i)$ representing average value for hyperbox H_i. This inclusion function is of general purpose type and finds similarity for data represented as multidimensional hyperboxes.

The degree of feature blocks inclusion represented by feature overlap function is defined as

$$v(F_i, F_j) = \begin{cases} 1 & \text{if } a(F_i) \in F_j \\ 0.5 & \text{if } F_i \cap F_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

with $a(F_i)$ representing average value for feature block F_i. This type of inclusion function leads to finding feature block similarity. Given feature block F_i the set of feature blocks similar to F_i can be determined.

The degree of data blocks and feature blocks inclusion represented by overlap function is defined as

$$\nu(N_i,F_j) = \nu(H_i,F_j) = \begin{cases} 1 & \text{if } a(N_i) \in F_j \\ 0.5 & \text{if } H_i \cap F_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

with $a(N_i)$ representing average value for data block N_i. Feature blocks F_i and data blocks N_i represent some kind of universe objects as given in corollary 4, 1 and may be compared by inclusion function. The inclusion $\nu(N_i,F_j)$ gives information which feature blocks are similar to given data blocks.

The degree of feature blocks and data blocks inclusion represented by overlap function is defined as

$$\nu(F_i,N_j) = \nu(F_i,H_j) = \begin{cases} 1 & \text{if } a(F_i) \in H_j \\ 0.5 & \text{if } H_i \cap N_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

with $a(F_i)$ representing average value for feature block F_i. Feature blocks F_i and data blocks N_i represent some kind of universe objects as given in corollary 4, 1 and may be compared by inclusion function. The inclusion $\nu(F_i,N_j)$ gives information which data blocks are similar to given feature blocks.

The degree of data blocks inclusion represented by overlap function is defined as

$$\nu(N_i,N_j) = \nu(H_i,H_j) = \begin{cases} 1 & \text{if } a(N_i) \in H_j \\ 0.5 & \text{if } H_i \cap H_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

with $a(N_i)$ representing average value for data block N_i. This type of inclusion function leads to finding data block similarity. Given data block N_i the set of data blocks similar to N_i can be determined.

In generalized deep covering approximation spaces, existence of universe objects coverings and feature coverings gives possibility of defining two additional rough approximation neighborhood systems.

Approximation neighborhood system F_S assigns two distinct feature coverings L, U for each universe objects coverings satisfying condition $L \subset U$ is called rough approximation neighborhood system F_S.

Definition 20 *Approximation neighborhood system F_S that assigns two distinct universe objects coverings L, U for each feature coverings is called feature approximation neighborhood system F_S.*

Definition 21 *Approximation neighborhood system F_S that assigns two distinct universe objects coverings L, U for each feature coverings satisfying condition $L \subset U$ is called rough feature approximation neighborhood system F_S.*

The concept of rough feature approximation neighborhood system has been defined and applied into generalized deep covering approximation spaces during construction of rough classifier framework. In the introduced rough extended model the universe with data represented by universe objects, is divided into data blocks N_i being covering of universe objects, features space is divided into feature blocks F_i. By selecting inclusion function of type $v(N_i, F_j)$ similarity of data blocks and feature blocks is determined. For each data block N_i its inclusion $v(N_i, F_j)$ in feature block F_j is calculated. Each data block N_i is described by its hyperbox determined by means of min, max feature block f_i values. Image blocks that have feature mean value in the F_i are assigned to its lower approximations, each feature block that is contained in the (min, max) hyperbox are assigned to its upper approximation. Given selected feature block F_j, lower and upper approximation of the feature block is defined as follows

$$G_*(F_j) = \{N_i \in U : \upsilon(N_i, F_j) = 1\},$$
$$G^*(F_j) = \{N_i \in U : \upsilon(N_i, F_j) > 0\}.$$

For the feature covering $F = \{F_1, \dots, F_s\}$ the lower and upper approximaions are defined as follows

$$G_*(F) = \bigcup_{F_i} \{N_i \in U : \upsilon(N_i, F_j) = 1\},$$
$$G^*(F) = \bigcup_{F_i} \{N_i \in U : \upsilon(N_i, F_j) > 0\}.$$

Introduced concept of generalized covering approximation spaces gives theoretical foundations into creation of rough feature covering model described in the next section. Metric spaces define the distance function that makes possible to compare objects, their similarity, relations, data structure and extract fuzzy and probabilistic properties of the objects of the universe giving at the same time interoperability of different data models.

3. Covering models for generalized covering approximation spaces

Covering models are created by means of data object relation to the selected set of data centers. This reference set of data objects performs as the set of thresholds or the set of cluster centers. In general data object are analyzed by means of their relation to the selected number of cluster centers. Cluster centers are regarded as representatives of the clusters.

Standard rough coverings - In standard setting data similarity is measured by means metric. In rough clustering approaches, data points closest to the given cluster center or sufficiently close relative to the selected threshold type, are assigned to this cluster lower and upper approximations. The upper approximations are calculated in the specific, dependent upon threshold type and measure way presented in the subsequent paragraphs.

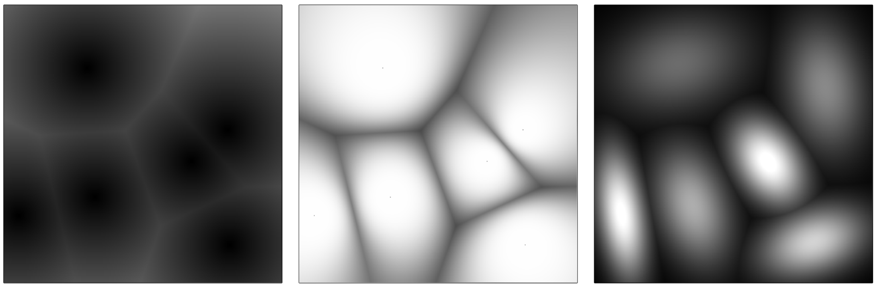
Fuzzy coverings - Fuzzy coverings are created by calculation of fuzzy membership values for all data points of the universe to cluster centers. Assignment to fuzzy coverings is based upon fuzzy membership values satisfying predefined conditions based upon threshold values.

Probabilistic coverings - Probabilistic coverings are calculated by means of probability distribution values within defined limits. In this type of covering probability distributions determines which data belongs to the coverings.

Neighborhood type - Neighborhoods are defined on the base of threshold value for the given type of coverings. Threshold neighborhoods defined data objects that are within given threshold value in the selected covering type, it means standard, fuzzy and probabilistic. Difference neighborhood define similar objects on the base their distance not higher to the nearest neighborhood.

Upper and lower neighborhoods - The upper and lower neighborhood approximations are defined on the base neighborhood type and feature type. In this way, lower approximation and upper approximation are defined independently.

In Fig. 1 three types of coverings have been presented. In Fig. 1 a standard coverings based upon standard metric has been presented. In Fig. 1 b fuzzy coverings based upon fuzzy membership values have been presented. In Fig. 1 c probabilistic coverings are presented.



a b c

Fig. 1. Three types of coverings, (a) standard coverings (b) fuzzy coverings, (c) probabilistic coverings

4. Rough covering classifiers in generalized covering approximation spaces

Construction of rough classifier is presented for image data but the classifier is applicable to any numeric data. The image is denoted as I=$\{p_1,\ldots,p_m\}$ representing image pixels and at the same time forming universe object according to definition 1. Each pixel forms n-dimensional point as values of arbitrary n features denoted as x=$\{x_1,\ldots,x_n\}$. In this way, each image consists of m pixels described by numeric features a. In next step, image blocks are defined as N=$\{N_1,\ldots,N_r\}$ with $K = \{1\ldots,r\}$ denoting the index set for image blocks and $N_i \subset I$ with $N_i = \{p_1,\ldots,p_r\}$. Image blocks are created according to selected block creation strategy. Feature space is divided into feature blocks F=$\{F_1,\ldots,F_s\}$ with $L = \{1\ldots,s\}$ denoting the index set for feature blocks. Further, for each image block $N_i = \{x_1,\ldots,x_r\}$ its min, mean, max values are calculated and hyperbox H_i is determined. Both image blocks and feature blocks may be distinct (separable) or may overlap.

For feature blocks F_k, image blocks N_i that intersect the hyperbox H_i are assigned to their upper approximations $F^*(k)$ and their measures $D^*(k)$ are increased by 1.0. The feature block(s) F_k that contain(s) mean value of N_i, is considered to contain entirely this image block so belongs to its lower approximation $F_*(k)$ and its measure $D_*(k)$ is increased by 1.0. The degree of data blocks and feature blocks inclusion represented by overlap function is defined as

$$v(N_i,F_j) = v(H_i,F_j) = \begin{cases} 1 & \text{if } a(N_i) \in F_j \\ 0.5 & \text{if } H_i \cap F_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

with $a(N_i)$ representing mean value for data block N_i. In case of precisely selected sets N_i and F_j only indices are used giving the following formula

$$d(i,j) = \begin{cases} 1 & \text{if } a(N_i) \in F_j \\ 0.5 & \text{if } H_i \cap F_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Lower approximation $F_*(i)$ of feature block F_i consists of all image blocks N_j that inclusion $d(i,j)$ is equal 1. The cardinality of all $d(i,j) = 1$ of feature block F_i is denoted as $D_*(i)$ and represents measure of the lower approximation

$$D_*(i) = card\{k \in K : d(i,k) = 1\}$$

Upper approximation $F^*(i)$ of feature block F_i consists of all image blocks N_j that inclusion $d(i,j)$ is greater than 0. The cardinality of all $d(i,j) > 0$ of feature block F_i is denoted as $D^*(i)$ and represents measure of the upper approximation

$$D^*(i) = card\{k \in K : d(i,k) > 0\}$$

Quality of approximation $QA(i)$ describes quantitatively ratio between lower and upper approximations of the feature block F_i

$$QA(i) = D_*(i)/D^*(i)$$

Roughness $R(i)$ of feature block F_i gives measure of the uncertainty of the feature block

$$R(i) = 1 - QA(i) = 1 - D_*(i)/D^*(i)$$

As a whole, all $R(i)$ roughness values describe rough descriptor of the universe data.

$$R = \sum_{i \in L} R(i) = \sum_{i \in L} (1 - QA(i))$$

Following presented procedure, the rough covering descriptor is calculated. The algorithm for classification based upon introduced descriptor has been presented in Algorithm 1. In the first step, all required data structures are prepared, it means input data are divided into data blocks, feature space is divided into feature blocks. In the second step, feature covering descriptor is calculated by assigning data blocks to feature blocks and to their lower and upper approximations. After feature covering descriptor has been calculated, the selected classifier is trained and tested.

Algorithm 1: Rough covering classifier

Prepare data structures
Calculate feature covering descriptor
Perform learning phase of selected classifier algorithm
Perform classification

Data structure preparation

Data block creation - data objects can be divided in arbitrary way that is called block creation strategy. For image data, data block creation strategy may consists in creation overlapping or non-overlapping image blocks. For feature block creation, the strategy described in previous Section may be applied with selection of cluster centers - cc, neighborhood type - nt and covering type - ct.

Algorithm 2: Preparation of data structures

Input - I - image, S - block creation strategy, T - covering creation strategy
Output - N = $\{N_1, \ldots, N_r\}$, F = $\{F_1, \ldots, F_s\}$
Divide image into image blocks $->$ N
Divide feature space into feature blocks $->$ F

Rough covering descriptor

Feature covering descriptor is created as for each data block all feature blocks that intersect bounding box of F_i and feature block that contains $a(N_i)$ are found and their lower and upper approximations measures are increased. The calculation of the feature covering descriptor has been presented in Algorithm 2.

Perform classification

After rough covering descriptor has been calculated for all data, classification may be performed by learning phase and testing phase. Rough covering descriptor may be used for any classification framework suited to this purpose such as Support Vector Machines, neural networks.

Algorithm 3: Rough covering descriptor

Input - I - image, N $->$ data blocks, F $->$ feature blocks
Output - R - quality of approximations and roughness of feature blocks F
foreach *Image block* N_i *in* N **do**
  Calculate hyperbox H_i and mean value a(N_i)
  foreach *feature block* F_p *in* F *which contains average value* $a(N_i)$ **do**
    Increment $D_*(p)$ of this feature block by 1.0
  foreach *feature block* F_p *in* F *contained in* H_i **do**
    Increment $D^*(p)$ of this feature block by 1.0

Computational feasibility

The most difficult part in calculation of rough covering descriptor is finding of all coverings embedded in hyperbox H. The routines Contains (F_i,N_i), Contains (N_i,F_i), Get(F_i), Get(N_i) may be helpful in accomplishing this task. Get(N_i) returns all feature blocks F_i that are intersected by hyperblock H_i. Get(N_i) works by calling for each feature block F_i function Contains(N_i, F_i) returning information whether bounding block of N_i intersects F_i. Function Get(F_i) returns all data blocks N_i that intersect feature block F_i. In case of the number of feature blocks is not large, all the feature blocks may be checked. When the number of feature blocks is large, for example feature blocks are generated from 1000 cluster centers, the routine Get(F_i) is more appropriate.

Conclusions and Future Research

In the paper, new algorithmic approach to construction of rough classifiers has been presented by introduction of rough covering descriptors. Introduced rough covering descriptor merges two different approaches in rough sets. Universe objects belong to coverings as covering approximation spaces. In generalized covering approximation spaces functionality of inclusion function has been given from generalized approximation spaces. Granularity concept is employed by selection of different types of data blocks and different feature blocks. Data descriptors are embedded in generalized covering approximation space with coverings in feature space and inclusion function that measures similarity of data blocks and feature coverings. Rough covering descriptors are based upon assignment of data blocks to feature blocks acting as data descriptor. In this way created rough covering descriptor can be easily used in classification frameworks such as Support Vector Machines framework, neural networks. Coverings in generalized covering approximation spaces include standard, fuzzy and probabilistic coverings giving robust theoretical framework in design, implementation and application of classification algorithms.

References

[1] Sai Y. Liu G.L. A comparison of two types of rough sets induced by coverings. *Int. J. Approx. Reason*, 50:521–528, 2009.

[2] Zhu W. Relationship between generalized rough sets based on binary relation and covering. *Information Sciences*, 179:210–225, 2009.

[3] Yao B. Yao Y. Covering based rough set approximations. *Information Sciences*, 200:91–107, 2012.

[4] Gomez J. Restrepo M., Cornelis C. Duality, conjugacy and adjointness of approximation operators in covering-based rough sets. *International Journal of Approximate Reasoning*, 55:469–485, 2014.

[5] Zhao Z. On some types of covering rough sets from topological points of view. *International Journal of Approximate Reasoning*, 68:1–14, 2016.

[6] Liwen Ma. Two fuzzy covering rough set models and their generalizations over fuzzy lattices. *Fuzzy Sets and Systems*, pages –, 2015.

[7] Skowron A. Pawlak Z. Rough sets: some extensions. In *Information Sciences 177*, pages 28–40, 2007.

[8] Stepaniuk J. Malyszko D. Granular multilevel rough entropy thresholding in 2D domain. *16th International Conference Intelligent Information Systems, Zakopane, Poland*, pages 151–160, 2008.

[9] Stepaniuk J. Malyszko D. Adaptive rough entropy clustering algorithms in image segmentation. *Fundamenta Informaticae*, 98(2-3):199–231, 2010.

[10] Stepaniuk J. Malyszko D. Adaptive multilevel rough entropy evolutionary thresholding. *Information Sciences*, 180(7):1138–1158, 2010.

[11] Peters J. Pal S. K. *Rough Fuzzy Image Analysis: Foundations and Methodologies*. CRC Press Inc, 2010.

Dariusz Małyszko

KLASYFIKATORY W UOGÓLNIONYCH APROKSYMACYJNYCH PRZESTRZENIACH POKRYĆ

Streszczenie W pracy przedstawiono nowy sposób konstrukcji klasyfikatorów w uogólnionych aproksymacyjnych przestrzeniach pokryć, definiowanych jako przestrzenie aproksymacyjne zawierające przestrzeń obiektów, pokrycia w tej przestrzeni, oraz pokrycia w przestrzeni atrybutów wraz z zdefiniowaną funkcją zawierania się zbiorów zastosowaną dla pokryć.

Słowa kluczowe: przestrzenie aproksymacyjne, przestrzenie pokryć, przybliżone przestrzenie aproksymacyjne