Egyptian Journal of Pure and Applied Science

Meta-Mining rule for governing the drift ratio of medical image

Gaber A. El-sharawy, Eman K. Elsayed and Eman S. Zedan[*]

Department of Computer Science, Faculty of Science, Al-Azhar University, Cairo, Egypt

ARTICLE INFO

ABSTRACT

In the database, the instance might change along with time; this problem knows concept drift. When the change occurs, the classification model builds using the old dataset is not suitable for predicting a new dataset. The classification of concept drift has attracted a lot of attention in recent years. Although many techniques have been proposed to solve this problem, they have been able to focus only on updating the classification model. The contribution of this paper, propose an association classification (AC) to achieve model more accurate and faster than those obtained with the traditional classification approach. AC is a new classification technique based on association rules in features selection function and evaluation process. From the previous technique, propose a meta-mining rule mining (2M-Rules) to discover the rules for governing concept drift.

1. Introduction

Data stream generates continuously arriving data with high speed and massive to applied in many real applications such as network intrusion detection and web click stream. Data stream has attracted attention in different fields such as database systems, data mining and distributed system. The data mining stream has been an interesting research topic for years such as detecting changes, data stream classification and data stream clustering. Data stream classification is the process of applying static classification technique to the data stream in order to predict useful information and evolving data stream from a continuous system in online fashion.

Classification of data streams generated from the market plane such as stock market streaming information. The majority research challenges in data stream classification are represented in concept drift to treatment in this paper such as the results of mining model change over time, control high speed, and unbounded memory of the data stream. It is interested in three main dimensions: reduce error rate of model to increase classification accuracy, reduce amount of space for computer memory using classifier model require constant memory, and minimizes the time required for building classifier from training examples to predict by reducing the number of attribute and records. The propose solution in order to develop systems capable of performing a wide range of tasks in various area of

research. Our solution based on previous research [1] have proposed concept drift rule mining algorithm called (CDR) to discover rules for governing in the concept drift.

In this paper, gives a brief introduction on data stream mining and describes the motivation and objectives of the paper. Association classification (AC) discusses in this paper to develop the extraction rules that governing in concept drift. This approach proposes a feature extraction technique based on association rules that also used to improve the classification accuracy. Now, present the proposed algorithm that is called meta-mining rules technique (2M-Rules) for governing concept drift. The pruning rule technique develops to reduce the complexity time in the evaluation process of the rule set without changing the solution significantly. Before the end of paper, discusses the experimental results are applied to medical image. Finally, our conclusions and future research directions are explained. Several authors have studied the idea of implementing to many a classifier technique for concept drift in the following section.

2. Concept drift overview

Concept drift is an important research for any data analysis involving temporally ordered data. This paper focuses on technique to suggest an agent, which is automatically tracking the drifts over interval time. The majority of concept drifts to demonstrate the significance of changes in different domain. If concept drift is observed then suggest handling technique. This paper

* Corresponding author.
  E-mail address: eman_zedan2012@yahoo.com

develops an algorithm for monitoring, detecting, learning and governing capable of handling concept drift. It also develops computational to suitable existing concept drift in the absence of online data. The treatment for concept drift on dataset can be divided into vertical direction or horizontal direction to give deeper insight to describe the drift problem. The concept drifts which effect on accuracy of model can be represented in three aspects: The vertical representation to the drift problem that known "unstable instance" when occur increase or decrease in the number of instances. Thus, the solution interests with approaches that treatment the drift of instances and statistical properties. Horizontally, suggest profession technique that used to handle drift on the features of dataset this type is known "Unstable attribute". Unstable attribute value is occurring when change of data due to the passage of time is called concept drift.

Addressing this problem could contribute significantly not only to detect, but also to the more general field of concept drift governing. To evaluate the aspects of this problem, i.e. the currently proposed solutions for concept drift problems, can be viewed as a sequence of datasets that discussed in this paper to improve the aspects of the drift problems. Nevertheless, this problem may be used for real application such as intrusion detection. Another way to introduce an expert system to the decision maker for governing in the drift process and extracting the rules to predict decisions which are very difficult task.

3. Literature review
3.1. Background and related work

Concept drift occurs when the model builds using the old dataset is not suitable for predicting a new dataset. Although old data can affect on the accuracy of classification models with negative, it is important keep track of this data in the current model. Some other challenges in machine learning based drift detection have been addressed in this paper. Historically, present the classifiers that uses to handle concept drift can be divided into many types: instance selection, instance weighting, ensemble learning, and incremental learning.

Instance selection: the goal is select instances relevant to the current concept. Example of instance selection such as window-based algorithm includes use window of fixed or dynamic size to construct a classification model, while other use heuristics for adjusts the window size to know the accuracy of classification.

Instance weight: the ability of the same learning algorithm to process weight instance. The classifier can be weighted according to their age. Experiments that instance weight techniques handle concept drift worse than instance selection techniques [2].

Ensemble classifier: dividing data into sequential chunks of fixed size and building an ensemble on those chunks may be effective ensemble is more efficient than building a single model. Ensemble classifiers for

handling concept drift using multiple classifiers to construct a proper predictive model. Each class will be assigned weight voting which is used to eliminate classifiers or not. The accuracy of the output of many classifiers is very sensitive to concept drifts in the evolving streams [3].

Incremental learning is opposed to batch learning. Batch learning for handling concept drift is examining a large collection of instances at once and forming a single model. Basically, incremental model evolves and update a model to more efficient for the task of the handling concept drift. There are many types of incremental algorithm to handle concept drift: ID3 rebuild classifier using an integrated version of new and old datasets [1].

ID4 has information gain in every node and judges if the best splitting attribute is still the same, after the addition of new instances. If they are the same, the new instance continues going to the next node and are evaluated again. If they are not, the sub-tree root in this node is discarded and new sub tree in then built so that decision tree is different arrangement nodes. ID5 preserves the value of splitting criteria in each node and determines the best splitting attribute if the best splitting attribute change, ID5 will replace them instead of deleting the sub tree, is called pull-up [4].

3.2 Some research issues in concept drift

The concept drifting problem in stream data classification has been addressed by several authors. Yang et al. [3] have proposed a general framework for mining the drift. The proposed technique uses weighted classifier ensembles to mine data streams. They use real life data streams to test their algorithms and compare between the single classifier and classifier ensembles. The proposed technique combines multiple classifiers weighted by their expected prediction accuracy. The selection of a number of classifiers instead of using all is an option in the proposed framework without losing accuracy in the classification process. Last [5] has proposed an online classification system can adapt to concept drift. The proposed technique rebuilds the classification model that dynamically adjusts the size of the training window and the number of new examples and uses the error rate as a guide to concept drift. The system uses info-fuzzy techniques for model building similar to decision trees and information theory to calculate the window size. Domingos & Hulten [6] has VFDT is decision tree based on Hoeffding tree. It splits the tree using the current best attribute taking into consideration the number of examining data items used satisfied a statistical measure which is Hoeffing bound. The last version of CVFDT [4] presents concept drift adapting to high fast CVFDT which keeps the model from changing concepts. It implements by monitoring the quality of old search decision tree respect to a sliding window of data from the data stream when it detects the distribution of data is changing at a node. In their comprehensive literature review across these domains, offer a definition of technological innovation, which is

particularly useful in the current discussion. Also is the best to capture the essence of innovation as it applies to new product development. Within the operations domain, the main properties of this research are summarized in the following features: high accuracy and fast adoption to change, incremental to receive data and send the result in any time, low computational cost in space and time, and minimal number of parameters. The previous studies have only dealt with the situation in the changes of instance, whereas our study focuses on the situation in the changes of overall datasets. In this paper, we highlight the benefits of these techniques to discuss any limitations, which encounter in the drift problem. This algorithm is initially proposed to carry out diagnosis tasks, and then adapted to govern concept drift.

4. Classifier algorithm descriptions

From classifier perspective, propose meta-mining rules technique to find rules for governing in the changes of data on interval time. Coming up with a flexible and meaningful meta-mining model is the key here. Data with meta-mining classifier may prove crucial in discovery, which is of enormous value to the scientific community. This technique carries out on all datasets have the same instance on interval time is necessary to avoid losing information. In this application, compare between datasets, if the same then no drift in attribute value. If the difference then detect drift in attribute value and integrate parallel between them. In agreement with Tsai, et al. [1], we demonstrate the development of concept drift rule algorithm (CDR) is motivated by the need for governing the concept drift. The development of the algorithm in addressing two points: the first point: integrate between all datasets over interval time. The structure of integration process will discuss in the next section, which generates to the requirement of drift problem. The second point: develop pruning rule technique to reduce complexity time in the evaluation process of the rule set without changing the solution significantly.

4.1. Association classification technique

In recent years, association rule mining active field to generate and analyze data for building classifier. This paper focus on extract unsupervised pattern to extract a supervised classification rule such as kill's two birds with one stone, this is called association classification. Association discovery and classification rule mining are essential for data mining applications. Association mining aims to discover descriptive knowledge for building a classification model from database. Complementary refer to Association Classification (AC), which is a new approach that integrates association mining and classification into a single classifier system.

In supervised learning, the input values given to the classifier that classifies training data. The training data is composed of the input vector is characterized by the features, to create the output vector is also called a class

label. When the input is given to the learning system, the learning algorithm generates the rule set. Intelligent creatures usually attempt to understand their environment with a simplified mode, which is called a model. The basic approach of these cognitive systems is to observe their environment and identify similarities between objects and events from their environment. These similar objects are grouped into a class and rules are constructed form the objects of this class then the model is capable of predicting the behavior of its environment. The aim of building a model is to understand the environment and the prediction. If the environment changes with time, or if the model is built from a slightly different environment, the model must be modified through some kind of learning. This technique is called Meta-mining rules to extract classification rules based on association rules in the same system. The purpose of AC is to establish and analyze the input data and develop an accurate description model for exploring the boundaries among classes.

Association Classification (AC) is defined mining then mining as performs association analysis on a set of data to produce information function and then find classification rules. In this paper, propose a scientific method for finding ways to extract the rule which achieve higher accuracy and efficiency of data. There are several functions of association classification, which is important and applied successfully to several areas: including the discovery of commodity deal dependence, customer relationship management, and risk analysis.

AC is an analytical tool used to discover rules that relationship between the data. The proposed technique is able to produce classifiers contain rules which achieve higher accuracy and efficiency. AC is more easily understood by human and greater flexibility in handling unstructured data. The users of AC mining algorithm have many times in search of the optimal results. AC is used in the field of decision science which plays an important role in supporting business, optimization problem, and scientific decision making. AC is easy to convert into a set of equivalent rules and used for making decision in business or computer programming. The procedure of the classification rule discovery from data consists of two parts: the reduction of attributes and the rule discovery. The requirements of this model are translated into many processes such as data collection, processing, rule mining, modeling and communicating the models quickly to assist the decision makers. The association classification technique can be divided into many steps: retrieve association rule from the giving training data set, determine gain ratio using association rule for extracting feature, the classifier can be easily build and translated into predictive rules, and evaluate the accuracy of rules by association rule.

4.2. Association rule criterion

For database consists of many instances, each instance contains a set of items and each item is associated with another item is called Association rule. It is well-know

technique in data mining to detect interesting relationships between items in the database. This approach based on incremental association rule discovery is important for applications in which dataset comes in the form of data stream. Association rule mining is applied in many fields such as decision making, disease diagnosing, drift detection, market development, and schedule optimizing. An association rules like clustering are forms of unsupervised learning. Finding association rule can be seen as a simple search problem is guided using rule support and confidence. The goal of association rule is used to detect interesting relationships between items in the database. Support and confidence rules are used to describe the interestingness and association power between feature and target class that illustrate with the rules:

• The support S (x) of item set x is the number of database record which satisfy x.

• Supports (x→ y) =s (x∩y) of association rule is the number of records that satisfy both x and y.

• The confidence c (x→ y) = s (x∩y) /s (x) is the number of records which satisfying both x and y to the number of records which satisfy x.

In the present time, show how the classifier build using gain ratio that is depend on association rule, from here show the discriminating power of the association rules are used to build an accurate classifier. When building a decision tree from association rule is less complex from building a classifier from data examples and the method assigns attributes to the nodes using criteria depend on the properties of attributes.

4.3. Feature extraction using gain ratio

Feature selection is an active research topic in the community of machine learning and knowledge discovery in databases (KDD). It contributes to make the rule mining model more comprehensible to domain experts, improving the detection and diagnosis performance and robustness of the model, and reducing model training. Feature selection also known feature reduction that is a technique to select a subset of relevant features for building robust learning models. Feature (attribute) selection has been an active field of research topic to play an important role as preprocessing for classification. The main goal of feature selection for producing a higher classification accuracy and improve the performance of learning model. Feature selection techniques intend to remove noise filter to increase the classification accuracy. Some feature selection methods have been proposed in term selection function. Designed with gain ratio is used as criteria to select the best feature for each node. The development of the algorithm is motivated to improve the accuracy of the classifier. It implements by gain ratio is based on association rule (valuable knowledge) that introduce strong function depend on association between each feature and target class. In this paper, suggest a gain ratio (knowledge criterion) is a good indicator to measure strong the relation between condition features and target class.

Furthermore, select the features whose gain ratio is higher than the average gain ratio is filter technique based on mining technique.

One of the methods is based on computing gain ratio for every feature and removing the features is smaller than the average gain ratio as certain threshold [7]. After feature selection technique, we will start with built classifier to extract the rule. The conditions that have attribute-value in rule are ordered by maximizing gain ratio.

4.4. Generation stage (Model building)

Classifier technique is widely used to handle data streams with drifting concept in an incremental manner. This work is interested in three main dimensions: increase classification accuracy, reduce amount of space memory, and minimizes the time required for building classifier from training examples to predict. In the production process, we must reduce the number of attribute and records that need to search each attribute. Predication algorithm is applied to the physical data transform for valuable knowledge to study the performance and test effectiveness of knowledge.

The classification problem can be sorted into non-incremental learning and incremental learning. Incremental learning is important for applications in which the training dataset comes in the form of data stream. Incremental learning is becoming ever more important since the most of the information in our lives is presented in data stream consists of the data block. Any instance of applicable data, such as disease variation, weather forecasts, consumers shopping habits, or virus detection may change along with time is known as concept drift. The current solution to the problem of concept drifts focus on efficiently rebuilding classifiers to accurately predict new incoming data sets. Data stream classification is the process of applying classification techniques in order to predict useful information and evolving data stream from a continuous system in online fashion. The main properties of classifier learning are the following: high accuracy and fast adoption to change, incremental and online to receive data and send the result in any time, low computational cost in space and time, and minimal number of parameters. Several authors have studied the idea of implementing many a classifier technique for concept drift in the following section. A classifier technique is widely used to handle data streams with drifting concept in an incremental manner. The user, they might be more interested in the rules of concept drift.

For example, doctors desire to know the main causes of disease variation, scholars long for the rules of weather transition and sellers would like to find out the reasons why the consumers shopping habits change. The computation powers specifying goals of the computation that operate data for generating and testing program specification to the search space of possible algorithm and explains both the creativity and efficiency of rules

strategy in three stages. The philosophical theoretical and technical definitions concentrate on various aspects of the problem. This technique for finding and describing structured patterns in data helping to explain data and make predication from it. The current research employs many data mining techniques for this purpose. The famous techniques developed for predication includes: Bayesian classification, neural network, genetic algorithms, and decision tree. Having presented ways to characterize a wide range of predication process in this paper, we introduce decision tree to capture knowledge and interpret in many different ways. The production process uses values associated with the instances in the learning to drive properties about the data. The learning algorithm is a powerful approach to discover knowledge from records of attribute values relate with the features. Discover set of features for identifying a set of class and then uses the features to build decision tree. The algorithm examines all available features and selects the features are most effectively split the set of samples into subsets for improving the separation between output classes. Traditionally, use decision tree can be created with a graphics program or specialized software. One of the main goals of this paper was to attempt for finding a way to predict who shows more task persistence. Informally, decision trees are useful for focusing the discussion for make a decision. Programmatically, it can be used to assign the values that possibly lead to outcome decisions, which can be automated. The tree structure shows how one choice leads to the next and the use of branches indicates that each option is mutually exclusive. Each instance in the training dataset includes a set of attribute values and target class, each internal node denotes a test on an attribute, and branch represents on outcome of the test. The results of decision tree or a set of rules to represent paths from root to individual learn and the probability measures of a particular rule. Each path from the root to a leaf node from a rule and each leaf node are associated with a target class. Ideally, the results of decision tree are an intelligently generalized version, which easier to compare using different datasets with similar information. Structure decision tree model to classifier can be easily build and translated into predictive rules which are used for predictive target class and pruning of the predicate rules and evaluate accuracy. The main objectives of classifier are the following: the first goal is producing an accurate classifier and the second goal understands the predictive structure of the problem.

4.5 Integration process over interval times

In change system, multiple trees are created and compare to measure similarity of derived model. So, the integration process is a good addition for governing the state of system over interval time. Within the context of the classification process, integration system deal with integrating many separate datasets on sequence time using "→" to work together for a achieving a single attribute value to produce a single output. Integrate process between all datasets to single dataset for implementing the classification model on a single dataset instead of users might require build the classification model for each dataset. As mentioned previously, one appears to have applied current knowledge of association classifier to the drift problem. To mine the drift rules using a 2M-Rules algorithm after integrating process between all datasets from different times. Therefore, the users can clearly and quickly find the drafting rules between all datasets for governing concept drift. The previous studies of the drift computation are implemented online that introduce in relate work. In this paper, introduce offline compute to control high speed for getting the results more details. The next section illustrates integrate all datasets into signal datasets to implement the proposed algorithm. To solve the results of model change over time, so need to build and integrate into on classifier model. The integration technique computes parallel to solve unbounded memory problem and extract the rules for governing the decision making process. This process is a good step to introduce the solution to the results of mining model that change over time. Now,this algorithm is adapted for investigating the changes in the meaning of concept drift.

5. Problem statement

For database D, which is called the training set consist of multiple records $r_1, r_2, \ldots \ldots r_N$ where $r = <y,c>$, N is the total number of instances in dataset. Multiple records r have k dimension feature vector $<y_1, y_2, \ldots \ldots \ldots \ldots, y_k>$ (independent variable) and class label C (dependent variable). Each attribute has v splitting value $s_1, s_2, \ldots \ldots, s_v$ and class label C has L classes $c_1, c_2, \ldots \ldots c_l$. After time, the values of dataset are changed, so that we proposed algorithm to discover rules that governing in concept drift. many data block $p^{t1}, p^{t2}, \ldots \ldots p^{tend}$ are selected from the same dataset in different time $t_1, t_2, \ldots \ldots,$ and t_{end}. The value of attribute k to instance i in many data block are $p^{t1}_{ik}, p^{t2}_{ik}, \ldots \ldots, p^{tend}_{ik}$ are respectively. In this problem, explain how support, confidence and number of conditions in rules are found and used to assign rank of the rule. Let RS, RC are the support rule which is set 1 and the confidence rule is set to 50% as the default that used to select rules higher than the default. Let N_R, N_c are the total number of instances which achieve the rule and the number of instances with class label c. df is the default value for goal predicate. A quantitative analysis to extract association classification modeling was studied in the four steps, which can be observed in the algorithm 1. This algorithm is to describe the steps of extraction concept drift rule which includes: determine gain ratio based on association rule for each attribute, feature selection using average of gain ratio, build classifier, and extract rule classifier to pruning technique and evaluate process.

Input: p^q, p^w, df, RS, Rc

Output: CDR-See5 Tree

Procedure:

If   every value in p^{t1}_{ik} = every value in p^{t2}_{ik} =………………..= every value in p^{tend}_{ik}

Then concept drift is not occur and stop.

Else if   every value in p^{t1}_{ik} != every value in p^{t2}_{ik}!=…………!= every value in p^{tend}_{ik}

Then concept drift occur and we start to extract rule.

Integrate many data set into one dataset have single attribute value $P^{t1-t2-……tend}$.

That represent into single dataset $p^{t1}_{ik} \rightarrow p^{t2}_{ik} \rightarrow p^{t3}_{ik} \rightarrow …… \rightarrow p^{tend}_{ik}$

Select data x from $P^{t1-t2-……tend}$, be training set have class label.

For every instance in x from combined data $P^{t1-t2-……tend}$.

If combine data $P^{t1-t2-……tend}$ empty

   Return df

If combine data $P^{t1-t2-……tend}$ are all of the same class c then

Return c

Else

// The main steps of extract gain ratio based on association Rule for each attribute.

The first step: Find Association Rule

Scan once to training data x with column k

Count   the   total   number   of   instance   N   to   extract   three   arrays count[l],count[v][k],count[v][k][l]

Extract distinct classes in class label

Count each class to find support

$$s(c_l) = \frac{|c_l|}{N}$$

// support of every class in class label is result from divide array count[l] on the number of instance

For each attribute in training data x

Extract distinct splitting value

Count each splitting value to find support

$$s(s_v) = \frac{|s_v|}{N}$$

//support of every splitting value is result from divide array count[v][k] on the number of instance .

Count each splitting value with each class to find support

$$s(s_v \rightarrow c_l) = \frac{s(s_v \cap c_l)}{N}$$

// support of every splitting value with every class in class label is result from divide array count[v][k][l] on the number of instance

Extract confidence

$$c(s_v \rightarrow c_l) = \frac{s(s_v \rightarrow c_l)}{s(s_v)}$$

// confidence of every splitting value with every class in class label is result from divide array count[v][k][l] on array count[v][k].

Stage (1)

The second Step: Extracting Feature Selection that used to build classifier

determine information gain (information detector) and gain ratio(knowledge detector) using association rule

Calculate Entropy (informative detector) of condition attribute k which is

$$H_i(y_k) = \sum_{v=1}^{n} s(s_v) \log_2 \frac{1}{s(s_v)}$$

Calculate Entropy of data C which is

$$H(C) = \sum_{l=1}^{m} s(c_l) \log_2 \frac{1}{s(c_l)}$$

Determine condition entropy of x given the variable y_k

$$H(C/y_k) = \sum_{v=1}^{n} s(s_v) \left[\sum_{l=1}^{m} c(s_v \rightarrow c_l) \log_2 \frac{1}{c(s_v \rightarrow c_l)} \right]$$

OR

$$H_i(C/y_k) = \left[\sum_{v=1}^{n} \sum_{l=1}^{m} s(s_v \rightarrow c_l) \log_2 \frac{s(s_v)}{s(s_v \rightarrow c_l)} \right]$$

Determine entropy(informative detector) of attribute y_k

$$H(y_k) = \sum_{v=1}^{n} s(s_v) \log_2 \frac{1}{s(s_v)}$$

Determine Gain Ratio (knowledge detector)for every attribute

$$GR(y_k) = \frac{H(C) - H(C/y_k)}{H(y_k)}$$

Scan once to training data x with column k to determine gain ratio as view in algorithm1
// Extracting Feature using average of gain ratio as threshold
Calculate the average Gain Ratio for attributes

$$AverageGR = \frac{\sum GR \text{ for every attribute}}{Number \text{ of attribute(k)}}.$$

Select distinct attribute which GR (y_k) > AverageGR and rearrange attributes with high gain ratio to produce concept drift rule.

For every instance in x from combined dataset.

If combine dataset empty

 Return df

If combine dataset are all of the same class c then

 Return c

Else

// The main steps of extract gain ratio based on association Rule for each attribute.

For each node

If node is pure or (can not be split) then

While all instances in node have common attribute value and the attribute is never selected as splitting attribute in the path from node to the root

Do

Go on splitting node by using attribute

Mark the path from leaf nod into root as rules

Stage (2)

The third step: Assigning rank for each rule

Assign the target class label by majority vote

Rule support $RS_0 = N_R$, rule confidence $RC_0 = (100\ N_c/N_R)$ %, and N_{RC} is number of condition for rule

calculate degree match between them using hammer equation

$$\text{hammer equation} = \sum_1^0 \delta$$

If $\delta = 1$ then condition attribute of rule R= attribute value of example E

Else if $\delta = 0$ then condition attribute of rule R != attribute value of example E

The percentage of label for target class (confidence) = hammer equation that satisfy rules / the whole number of condition in rules that represent rule confidence.

$$\text{Accuracy of classifier} = \frac{\text{sum all confidence of rules}}{\text{the number of rules}}$$

Stage (3)

Algorithm 1: Association Classification Description in three stages

If we want camper between two rules r_1, r_2 in rank and we say rule r_1 is higher ranked than r_2

1. If r_1 has higher confidence RS_0 than r_2.
2. If both rules r_1, r_2 has confidence (RS_0) are equal, then r_1 have support (RC_0) is greater than r_2.
3. If two rules r_1, r_2 are equal in confidence (RS_0) and support (RC_0) then r_1 has less condition (N_{RC}) in left hand side than r_2.
4. If two rules r_1, r_2 are equal in confidence (RS_0), support (RC_0) and number of condition (N_{RC}) on the left hand side to rule then r_1 was generated earlier than r_2.

After assigning rank of a rule, we are resorting rules with higher rank to facilities apply pruning technique and evaluate algorithms. The rule ranking computes by of association rule which is an important step to help in choosing the most effective rules for prediction. Partially, they include confidence "accuracy", support, and the number of conditions for the rules as shown in this algorithm 2.

5. Rule Pruning

One way to improve the classifier accuracy of decision trees is pruning rules. It plays an essential role in designing classifier whereas decides the rules belong to final the classifier because pruning allows a search technique to be completed in reasonable time. In this context, we show the quality of rule sets from the standard algorithm1 for governing in object tracking to implement pruning rule. For classification purposes quickly, it is recommended to reduce the complexity of the rules set without changing the solution significantly and the result is better with optional pruning. A new technique provides efficiently to decrease the error rate, reduce the size of the rule set without losing its discriminating power and improve the generalization performance of rule classification.

The key motivations to pruning rules are: reducing the complexity of the rules and improve the accuracy of rules classification. Pruning technique introduces the contribution for the saving of the main memory, and reduces time and the number of searching operations in the evaluation process and the search cost. The rule pruning algorithm can be listed in the following steps:

After building classifier and convert into set rule, we calculate rank for each rule in the classifier.

1. For every rule in rule set of the classifier
2. Pruning condition by drop any single condition in set of rules. We start with two rules r_1 and r_2.
3. If r_1 and r_2 have attribute values in the conditions and class label are the same but different in attribute value of final condition Then merge rules into single rule have a class label with the highest rank (remove rule with lowest rank).
4. If r_1 and r_2 have attribute values in the conditions and class label are the same but different in a number of conditions Then remove the rule has fewer in number of conditions.
5. If r_1 and r_2 have attribute values in the conditions are the same but class label are different Then merge two rules into single rule have a class label with highest rank remove rule with lowest rank.
6. If r_1 and r_2 have attribute values of the conditions are the same but different in values of class label and number of conditions Then append new condition have attribute value which different about other rule to rules and have fewer condition.
7. Repeat steps 4-7 until no rules can be merged.

6. Evaluation based on computation association rule

In this paper, the accuracy of classifying determines using association rules to evaluate that used as the filter method to the rules. It is used to build structures and evaluate the classifier using rule ranking that called "filter rules". The rule ranking computes by of association rule, which is an important step to help in choosing the most effective rules for prediction. Partially, they include confidence "accuracy", support, and the number of conditions for the rules can be identified from this algorithm 2. We introduce an evaluation algorithm using rule ranking, which is the most important computation in testing data and deployment data. The previous studies of algorithms used to convert low level features into high level diagnosis.

The fourth step: Evaluate the data using rule ranking

Input: R (rule), E (example)

Output: The target class and accuracy of classifier

Procedure:

For each example E in testing data

Insert example E to model of rule classifier

Scan to every rule classifier

If all rules have the same class then predict the same class for example

Else if there are no rules active for example then return default class

Else compare between condition attribute of rule R and attribute values of example E by ranking rule

If there are two rules r_p and r_q predict different classes for the example and r_p is higher rank than r_q

Then select rule r_p for the example and retrieves target class of rule r_p.

End for

Algorithm 2: Evaluate rules using rule ranking

7. Experimental Results

In the experimental environment, the proposed algorithms are done on personal computer with Windows 7 professional. 2M-Rules algorithm run in Microsoft visual C# for extracting knowledge and performance evaluation. The bio-medical image system is used to illustrate the important role of technique in the real world. In this work, we implement 2M-Rules the algorithm on the patients of breast cancer. In image preparation stage, digital analysis was generated by JVC.TK.1070U color video camera mounted on top of an Olympus Microscope and the image was projected into the camera with a 63×objective and a 2.5×ocular. The first step of image analysis, image was captured by computer Eyes/RT color frame grabber board as a 640×400 Targa file. A bit-per-pixel grayscale image is used for the image analysis, which is not interested in the color of the nuclei. The exact location of each cell nucleus is specified and developed by a graphical computer program called Xcyt that allows the user to input the approximate location of enough nuclei (10-20) to provide a representative sample [8].

For monitoring of the images are displayed on a computer monitor can be enhanced before they are printed on the film or insert into segmentation software. The partitioning process of image that called "segmentation process" as shown in figure1, which divide image into multiple segments in three stages. It is to identify and segment into some meaningful objects according to the feature of the image. Images can be adjusted by radiologist to improve the accuracy of diagnosis may reduce the number of follow up procedures.

In our experiments, the breast cancer diagnosis dataset contains on target class to specific the type of breast cancer, which depend on analysis into ten basic features in 3-dimension. Every dataset includes instances that represent the objects of one image are applied to experiments as training dataset.

The possibility of data sets generated by generator software ware has several well-defined to generate several kinds of datasets that useful to evaluate our meta-mining rules techniques. The generation datasets that extract from the generation process of the image contents on interval time to evaluate the proposed technique for extracting the results. The objects of the image are represented as instances that applied in experiment as training dataset. After the diagnosis process using meta-mining rules technique then specific evaluation technique to compute the percentage of label for a target class with confidence rules.

In one image, the computation of object separate on each other of the object in the image that include the percentage and accuracy of label that useful in computing drift ratio. In diagnostic process, the percentage of label for target class is computed to help the decision maker in determining drift ratio of the image. The testing instances of objects that really have drifted concept are used to calculate accuracy. We demonstrate how the accuracy of the algorithm is affected by different drift levels and compare between them. It contains drift ratio each object is equal to the difference between the current and previous of percentage of drift point between images. The accuracy of the diagnosis process represents on dimension y in figure 2 to illustrate the images of fours patients. This figure is the most important of experimental results have been shown by our algorithms. These algorithms are able to view all drift ratios of the image are indicated in figure 2.
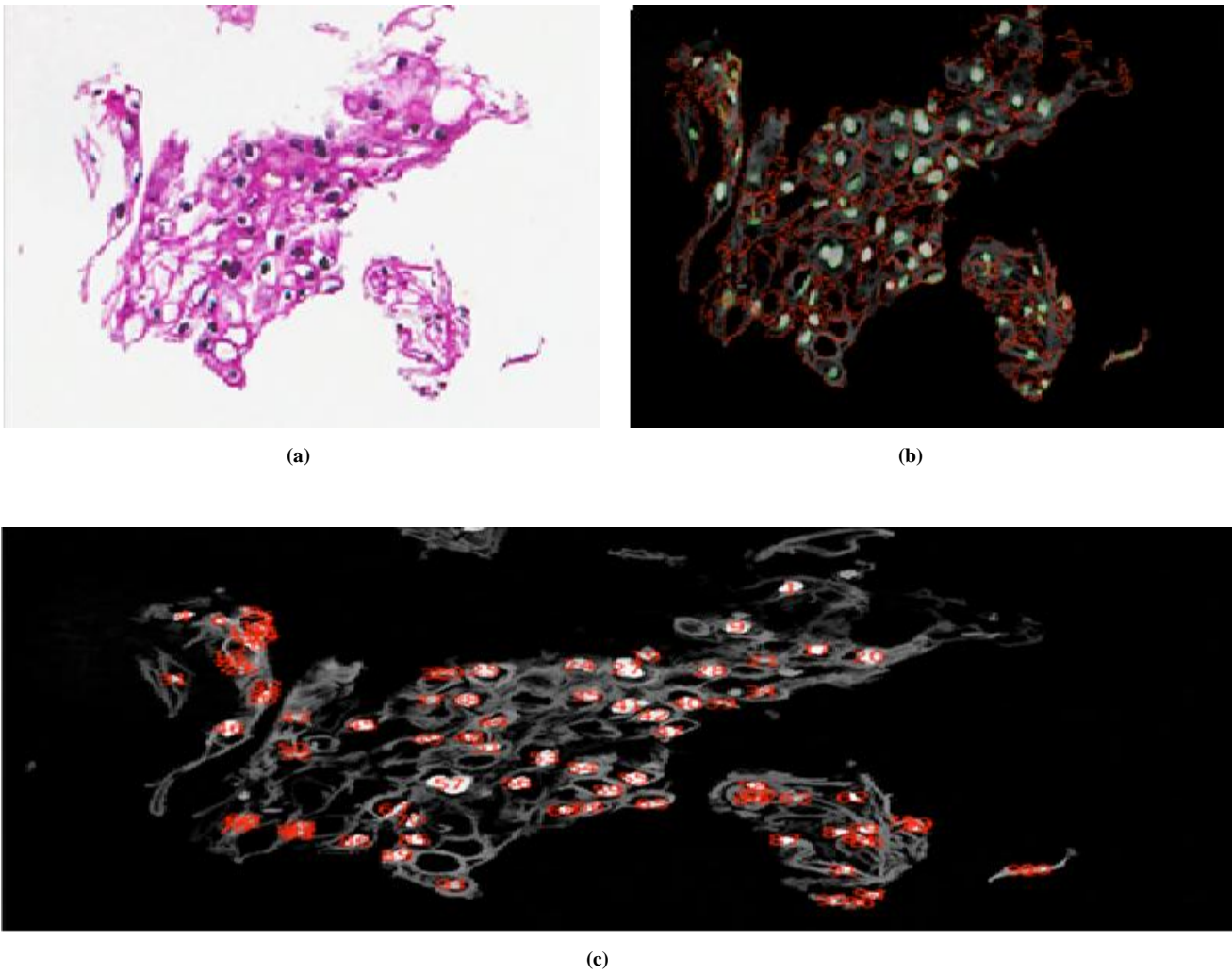
(a)

(b)

(c)

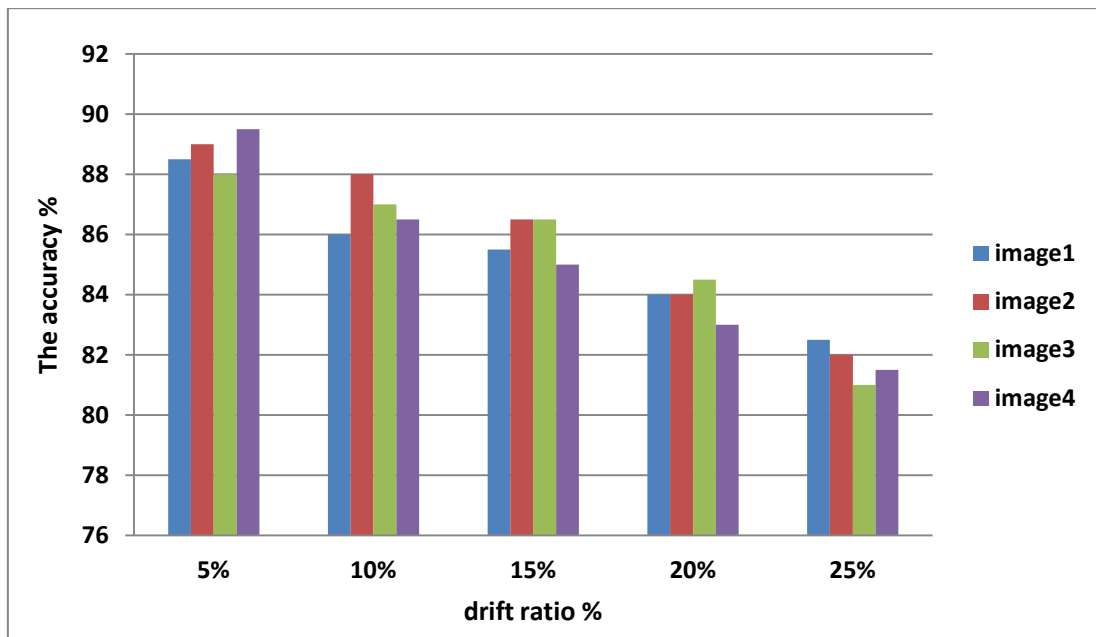**Fig. 1:** The representation for the stages of image segmentation.

Fig. 2: The accuracy of 2M-Rules algorithm under five different drifts ratios

From the previous results can be displayed in figure 2, we introduce knowledge in order to analyze the performance of our algorithm under different drift ratio. The proposed algorithm presents efficiency against the traditional techniques in the previous studies. It represents by comparison between the accuracy of the 2M-Rules algorithm of five different drift ratios for the four images of the patient. The main principle of our program, shown that the drift problem caused by variation of some feature. The integration process is used to integrate between data sets to extract from segmentation software into a single dataset. This is computed in algorithm to extract the rules for governing in the state of the patient over interval time. This paper focuses on compute the accuracy of diagnosis for the image can be found in the previous figure. The concept drift rules produce by the 2M-Rules algorithm can accurately predict the drifting instances. For each experimental data set, we only select the instances that really have a drifting concept from the testing data to calculate the accuracy. The concept drift rules mined by the 2M-Rules algorithm can accurately predict those drifting instances. The motivations behind the 2M-Rules algorithm for providing rules and quickly extract the prediction model if it is required by users.

8. Conclusion and Future Work

The user might be more interested in the rules of concept drift. The most important reason of use association rules to detect drift early to avoid the remainder of computation classifier. In this study, we address this issue and propose an algorithm based on association rule that is used in splitting function. The goal of this algorithm is to introduce splitting function based on mining technique (valuable knowledge) to select features that is used to build the classification model. The first advance of knowledge criterion for detects drift of the features over interval times that depend on the target class. This process is called multiple mining that means mining to extract association degree of the features and then mining for limiting the target class. The 2M-Rules algorithm cannot only produce the drifting rules, but also the classification model of each image for decision makers to have wide application. The evaluation process using association rules to label the percentage of type of cancer that useful to identify drift ratio of the object. The goal of this paper, the model is used to label each object in image by the feature that extract from segmentation software. This algorithm more efficient and useful process to image have drift problem. In our experiment, the decision makers have proposed different methods, that all focus on extract classification rules of the system. The pruning rules are used to reduce the number and size of rules to reach the results in suitable time.

After pruning, the classification rules are implemented in the medical system to diagnose the types of disease. The accuracy of the algorithm is inverse proportion to drift ratio of the images. In this paper, the integration process is used to integrate datasets of the image and extract rules that governing of the state of the patient at interval time.

In the Future we will focus on introducing the methodology to represent extraction knowledge from classification model that most useful and suitable in a medical information system. We introduce many computations intelligent methods to compute drift ratio and extract high or low the drift point in the image. In the Future focus on enhancing and applying this technique on medical image for illustrating important the algorithm in clinical practice.

References

1) **Tsai, C., Lee, C. and Yang, W. (2009).** Mining decision rules on data streams in the presence of concept drifts. Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan.

2) **Kolter, J.Z. and Maloof, M.A. (2003).** Dynamic weighted majority. A new ensemble method for tracking concept drift. In the processing of the third international IEEE conference in data mining (pp.123-130).

3) **Yang, W., Wang, H. and Han. J. (2003).** Mining concept-drifting data streams using ensemble classifiers. In Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining. Washington, DC.

4) **Hulten, G., Spencer, L. and Domingos, P. (2001).** Mining time-changing data streams. In Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining (pp. 97–106). San Francisco.

5) **Last, M. (2002).** Online classification of nonstationary data streams. Intelligent Data Analysis, **6(2)**: 129-147.

6) **Domingos, P. and Hulten, G. (2000).** Mining high-speed data streams. In Proceedings of the sixth international conference on knowledge discovery and data mining (pp. 71–80). Boston.

7) **Maass, W., Uchizawa, K. and Douglas, R. (2006).** Energy Complexity and Entropy of Threshold Circuits. International Colloquium on Automata, Languages and Programming,(CALP), Italy.

8) **Wolberg, W.H., Street, W.N., Heisey, D.M. and Mangasarian, O.L. (1995).** Computerized diagnosis of breast Fine-Needle aspirates. Departments of Surgery and Computer Sciences, University of Wisconsin, Madison, U.S.A.