TEXT MINING CLASSIFICATION APPROACHES FOR ASSOCIATION RULE MINING SET VALUED ATTRIBUTES

Mr. P. Pavankumar[1], Dr. Rashmi Agarwal[2]
[1] Research Scholar, CSE Dept, Madhav University, Abu Road, Sirohi, Rajasthan.
[2] Associate Professor, Computer Science and Engineering, Madhav University, Abu Road, Sirohi, Rajasthan.

Abstract— Association Rule Mining (ARM) and classification are integrated together to build competitive classifier models called Associative Classifiers and this approach is known as Swarm Intelligence (SI) and Ant Colony Optimization (ACO) SI leads to the formation of accurate classifier consisting of sub segments, the Ant Colony Optimization (ACO) meta-heuristic is propelled by the searching conduct of ants. These ants will probably locate the most brief developed by the ants speaks to a potential answer for the issue being tackled. We have be proposed such as Classification Based Association (CBA) rules, Post Classification Association, Association Rule based Classification Model(ARCM). These algorithms gives a Association Rule Mining with Set-Valued Attributes and Classification with Set-Valued Class Attribute, We have extended our classifier to have the capacity to handle set-valued class attributes and have developed novel techniques to predict set-valued arrangement.

Keywords: Association Rule Mining (ARM), Classification Based Association (CBA), Post Classification Association..

I. INTRODUCTION

1.1 Association Rule Mining

Association rule mining is one of the essential and all around investigated techniques of data mining to discovery it al connections among data things. In light of the design of data base, 28 distinct techniques have been created for mining the data. The flat design of database is utilized by Apriori Arrangement strategies while vertical format is the base of FP development and Éclat calculations. A different methodology either improves the productivity of the current methodologies or manages abnormal state data mining ideas.

The most agreeable order of data mining techniques is on the premise of the design of the database under thought. Diverse methodologies have been suggested that utilization even design of database, vertical format of database or anticipated format of database. A few scientists deal with enhancing the productivity of the mining process while others attempted to uncover progressed, confused and abnormal state information from the database. Additionally, swarm insight techniques have been utilized as a part of various fields for different assignments going from advancement to appropriation of assets. The utilization of swarm insight for data mining has turned out to be well known since most recent

two decades. After that few developments in the field of data mining utilizing swarm knowledge has been completed. This section contemplates light on the accessible writing in both the field's viz. data mining and swarm knowledge and likewise introduces a talk of the fruitful applications of various swarm insight techniques in data mining.

1.2 Swarm Intelligence

The two standards of swarm knowledge territory are: Ant Colony Optimization [DOR2004] and Particle Swarm Optimization [KEN1995]. Since most recent two decades these techniques have spread their impact in every aspect of enhancement. A few variations and strategies for these techniques have been contrived after some time. The well-known applications of these techniques are talked about in next sub-segments.

1.3 Ant Colony Optimization

The Ant Colony Optimization (ACO) meta-heuristic is propelled by the searching conduct of ants. The's ants will probably locate the most brief developed by the ants speaks to a potential answer for the issue being tackled. ACO has likewise been utilized as a part of applications, for example, rule extraction, Bayesian network structure learning, and weight advancement in neural network preparing.

Dorigo et al. [DOR1991] presented the term ACO and displayed a structure for the utilization of the strategy in various ranges. The particular prerequisites and general strides that must be followed for applying the ACO based drew closer were begat. Utilizing the manufacturing conduct of ants for tackling complex enhancement issues was exhibited. Ants convey by implication through adjusting the earth and utilize an input component to draw in different ants. As an ever increasing number of ants take after a trail, the odds of receiving the trail by all the more up and coming ants increments and this at last prompts joining of the arrangement. A similar thought is connected in simulated subterranean insect framework for getting ideal arrangement.

At last the underlying pheromone estimation of each trail is set at Gambardella and Dorigo [DOR1997a] exhibited Ant Colony System (ACS) that adventures the pursuit encounter collected by the ants more firmly than Ant System. Additionally pheromone vanishing and pheromone store happen just on the bends having a place with the best-so-far visit. The ants expel some pheromone from the bend to build the investigation of the option ways.

1.4 Particle Swarm Optimization

The PSO meta-heuristics is propelled by the facilitate development of fish schools and winged animal runs. The PSO is exacerbated by a swarm of particles. Every molecule speaks to a potential answer for the issue being illuminated and the position of a molecule is dictated by the arrangement it at present speaks to.

1.5 Other Swarm Intelligence Techniques

Yang, X.S. [YAN 2009] examined the conduct of Lampyridae bioluminescence. Lampyridae is a group of bugs that are proficient to deliver regular light (bioluminescence) to pull in a mate or a prey. They are ordinarily called fireflies or lightning bugs. The firefly calculation (FA) was connected to the streamlining of benchmark capacities.

Yang, X. [YAN2010] proposed a Bat Algorithm (BA), in view of the echolocation conduct of bats. It is conceivably more intense than molecule swarm enhancement and genetic calculations. The essential reason is that BA utilizes a decent mix of real favorable

II. EXISTING WORK

2.1 Data Mining with Ant Colony Optimization

Subterranean insect state improvement system has been as of late connected to data mining assignments. The most unmistakable utilization of ACO in data mining is in classification rule discovery. Grouping has additionally scope for utilization of ACO, particularly brood arranging procedure and numerous analysts connected these strategies for bunching assignments. However utilization of ACO for association rule mining is constrained. The most striking looks into have been talked about below:

2.2 Classification Rule Discovery with Ant Miner

Data mining comprises of a few errands and ACO has been connected for a large portion of these. Classification is a vital data mining errand, where the estimation of a discrete (subordinate) variable is anticipated in view of the estimations of a few autonomous factors. ACO inside the data mining group has been utilized basically for administered classification. In spite of the fact that ACO has been fundamentally utilized for grouping, the majority of the exploration tends to the territory of classification rules discovery. The most eminent research toward this path is the Ant Miner rule enlistment system. Ant Miner was the principal use of ACO to classification detailed by Parpinelli et al.

The Ant Miner, a successor of Ant Miner proposed by Martens et al. [MAR2007] contrasts from the Ant Miner form in a few ways. The earth is characterized as a coordinated non-cyclic diagram (DAG), with the goal that the ants can pick their ways all the more successfully when the Ant Miner condition is completely associated.

AntMiner2 [LIU2002] additionally expanded Ant Miner yet utilized a more straightforward, however less precise thickness estimation condition when contrasted with the heuristic esteem. This made AntMiner2 computationally more affordable without a critical debasement of the expressed execution.

AntMiner3 [LIU2003] likewise expanded Ant Miner by presenting two noteworthy changes, bringing about expanded precision. Right off the bat, an alternate refresh rule is utilized which is characterized with the nature of a rule set to the entirety of its affectability and specificity. Also, more investigation is energized by methods for an alternate move rule that builds the likelihood of picking terms not utilized as a part of already developed rules, as executed by the Ant Colony System.

2.3 Clustering with ACO

The data grouping calculation proposed by Lumer and Faieta called LF calculation depends on an indistinguishable subterranean insect conduct from displayed by Deneubourg et al. Increments to this essential LF calculation that have been proposed are a versatile setting of the parameters, enabling different data things to be transported on the double.

An enhanced variant of the LF calculation has been proposed, named ATTA (Adaptive Time-subordinate Transporter Ants), which has given its detailed outcomes and rather broad changes. ATTA has two variations: one which is restricted to a topographic mapping, named ATTA-M, and one which really brings about bunches of data (ATTA-C). The primary advantages of the ATTA-C system is an unequivocal dividing without need of human intercession and an earlier setting of parameters relying upon dataset characteristics

2.4 Association Rule mining with ACO

The creators have proposed Artificial Bee Colony Optimization calculation for concealing the Veenu Mangat has worked in the field of rule mining for therapeutic area. She has revealed that therapeutic space produces a colossal amount of data day by day. So removing valuable data and giving help to logical basic leadership, with the end goal of analysis and treatment of sickness, has moved toward becoming need. The creator has displayed different techniques for rule mining in restorative space, recognized holes and proposed a novel half and half system for productive rule mining [MAN2012]. In another paper, the creator has revealed the utilization of swarm insight for classification rules in therapeutic space and proposed a joined ACO/PSO approach for removing classification rules.

III Proposed Work

2.1 Association Rule Mining

2.1.1 Notations and Basic Concepts

Association rule mining is the technique of finding association rules that satisfy the predefined minimum support and confidence from a given database. This technique is widely adopted in the market basket analysis and currently used in various fields where relativity of the attributes plays a vital role in deciding the functionality of respective domains. Association rule is a relation between a pair of disjoint item sets. If LHS and RHS are two disjoint sets of items, the association rule is stated as LHS → RHS. LHS and RHS are sets of items, the RHS set being likely to occur whenever the LHS set occurs

2.1.2 *Formal Problem Description*

Let I = $\{i_1, i_2 \ldots i_n\}$ be a set of items. Let D be a set of transactions or database. Each transaction t€D is an item set such that t is a proper subset of I. A transaction t supports X, a set of items in I, if X is a proper subset of t. An association rule is an implication of the form $X \rightarrow Y$, where X and Y are subsets of I and $X \cap Y = Ø$. X is called the antecedent and Y is called the consequent part. The support of rule $X \rightarrow Y$ denotes ratio of the number of transactions in the database that contains the itemset X and Y to total number of the transactions in the database D. The confidence of rule is the ratio of the number of transactions in the database that contains the itemset X and Y to number of the transactions that contains X. A rule $X \rightarrow Y$ is strong if it reaches the minimum support threshold and minimum confidence threshold. Association rule mining algorithms scan the database of transactions and calculate the support and confidence of the rules and retrieve only those rules having support and confidence higher than the user specified minimum support and confidence threshold [MIN1996]. Association rule mining consists of two stages viz. the discovery of frequent itemsets and the generation of association rules. It follows that in majority of cases, the discovery of the frequent set dominates the performance of the whole process.

2.1.3 Frequent Item set Generation

Visit thing sets are those arrangements of things whose events surpass a predefined edge in the database. The computational prerequisites for visit itemset generation are generally more costly than those of run generation. Generating all the subsets of set of things, say I {i1, i2… in} for large value n, is practically a lot of troublesome because of the gigantic search space. In fact a linearly developing number of things suggest an exponential developing number of itemsets [ZAK1997]. The way toward generating incessant itemsets can be additionally isolated into two sub-issues: Candidate large itemsets generation handle and successive itemsets generation prepare. Itemsets that are normal or have the plan to be large or incessant are called candidate itemsets and among those the itemsets whose help surpasses the help edge are viewed as regular itemsets.

2.2 Classification based Association Rules (CBA)

The rules coming about because of Associative Classification mining can be assessed to choose a subset of the rules that will frame the model or classifier. To the best of our insight, Liu, Hsu, and Ma were the rest to create a classifier in light of affiliation rules. They demonstrate that the classifier constructed executes and also or superior to anything surely understood decision tree calculations. From that point forward, numerous affiliation administer based classifiers have been worked for different areas. Among others, for classifying mammography pictures, for classifying web documents, for recommender frameworks, for classifying spatial information, for document classification, and for content arrangement. The way toward building the classifier includes choosing rules by certainty or support. Certainty is a well known standard for run determination to the classifier as it signifies the strength of a run the show. On account of CBA, they utilize a heuristic to choose a subset of the rules that orders the preparation set generally precisely. At times, the pruning is as basic as evacuating contradicting rules or more confused like utilizing post pruning techniques that are utilized as a part of decision trees. In CBA - CB, the created CARs are requested in light of the accompanying definition.

Definition 2.2.1 Rule Ordering Association

Given two rules, ri and rj, ri, rj (ri goes before rj) if the certainty of ri is more noteworthy than that of rj or, their certainty are the same, however the help of ri is more prominent than that of rj , or, both the certainty and the help of ri and rj are the same, yet ri is produced sooner than rj .

Give R a chance to be the arrangement of CARs and D be the preparation information. The point of the model development algorithm is to pick an arrangement of profoundly predictive rules in R to cover the preparation information D. The classifier constructed is of the accompanying structure: $< r_1; r_2 ; ..... ; r_n;$ default class $>$ where r_i 2 R, $r_a r_b$ if $a < b$. Default class is the default mark utilized when none of the rules can classify a case. Algorithm 2 demonstrates the CBA-CB technique. In stage 1, the rules are sorted by the request said above; at that point each administer is considered thusly.

Algorithm 2 CBA-CB Algorithm
Inputs: rules R, training set instances D
Output: classifier C

1. R = sort(R);
2. for each rule r 2 R in sequence do
3.   temp = ;
4.   for each instance d 2 D do
5.   if d satisfies the conditions of r then
6.   store d.id in temp and mark r if it correctly classifies d;
7.   end if
8.   end for
9.   if r is marked then
10.   insert r at the end of C ;
11.   delete all the cases with the ids in temp from D;
12.   select the default class for the current C ;
13.   compute the total number of errors of C ;
14.   end if
15. end for
16. Find the r_{st} rule p in C such that C_p, the list of rules in C up to p, has the lowest total number of errors. and drop all the rules.
17. Add the default class associated with p to the end of C, and return C

3.1 Implementation

We have implemented our classification framework in WEKA. WEKA is an open-source suite of machine learning algorithms. The inspiration for actualizing our theory in WEKA is the broad utilization of this framework in WPI's Knowledge Discovery and Data Mining Research Group. WEKA is created in the Java Programming Language. Figure 4.1 demonstrates the engineering of our classification framework. We modified the current Apriori like algorithm, Apriori Sets and Sequences, to generate classification association rules. The generated rules are utilized for building models. The resulting models are tried for accuracy.

Alluding to Figure 3.1, the association rule based classification algorithm is called Associative Classification and is a piece of the Wpi. Classifiers bundle. We demonstrate the communication between Associative Classification and Apriori Sets and Sequences. We additionally demonstrate the distinctive modules in both the algorithms.

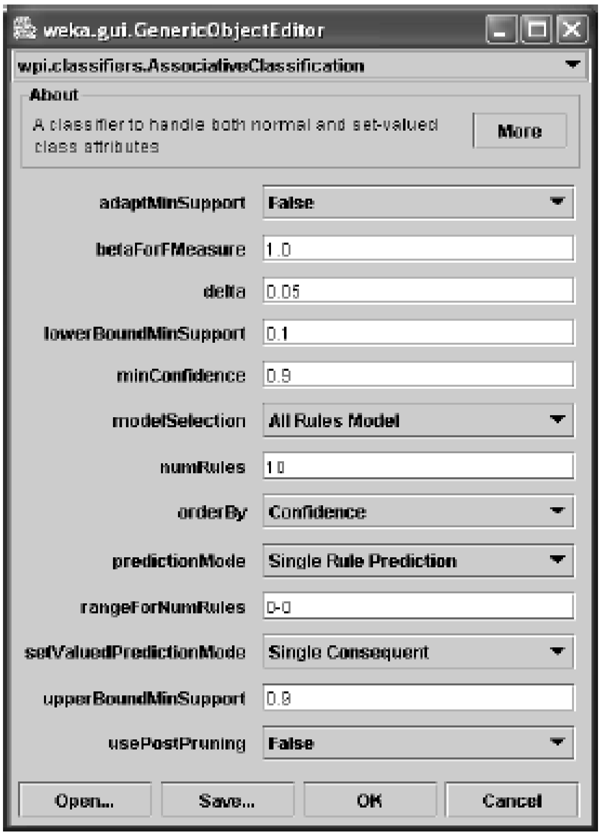

Figure 3.1: Parameter Menu for Our Extended Association Rule Mining

Algorithm 3 is the modified control procedure to mine (compelled) CARS. In this proposal, we have modified the first control procedure (see Algorithm 1) to permit pruning of rules in light of pessimistic error. We have additionally modified the algorithm to take into account nearness or nonappearance of items (semantic constraints) in the rules. All the more accurately, clients can specify an item to show up or not to show up on either the forerunner or the subsequent of a rule. We utilize a pruning technique to generate just item exhausts, threshold, sets that can potentially turn out to be a piece of the client indicated rules. In the rule generation advance, before a rule is generated, it is verified whether it fulfills the client determined constraints and, provided that this is true, the rule gets generated. WEKA contains numerous outstanding classification algorithms and one noteworthy commitment of this proposition is the classifier in light of Association rule mining algorithm.

Info parameters incorporate required Antecedent, required Consequent, dis-allowed Antecedent and dis-allowed Consequent. The while circle in Step 5 rehashes itself until the point when the support limit is beneath the min support or the quantity of rules generated success as per the client indicated number of rules. In the event that we investigate the iterative procedure of generating item sets and rules from them: In Step 6, we generate the 1-item item sets. In stages 11-15, the condition debilitates all conceivable item sets that can be created until the point when no more items of size k can be joined to deliver items of size (k+1). Just those item sets that will potentially yield rules with the required item sets are generated. The algorithm for this can be found in Section 4.3.

Algorithm 3 Modified Apriori Sets and Sequences Control Procedure

Inputs: required Antecedents, required Consequents, disallowed Antecedents, disallowed Consequents, num Rules

Outputs: rules

1. rules = ;
2. support = UpperBoundSupport;
3. freqItemsets = ;;
4. requiredItems = requiredAntecedents [required Consequents]
5. while (support > minsupport AND rules. Size < numRules) do
6. L_1 = f1- item itemset g;
7. for (k = 2; L_{k1} 6= ;) do
8. C_k = generate Candidates(L_{k1} , required Items);
9. L_k = evaluate Candidates(C_k);
10. FreqItemsets [L(k)];
11. end for
12. maxFreqItemsets= genMaxFreqItemset(freqItemsets);
13. rules = Generate All Rules(maxFreqItemsets, requiredAntecedents, requiredConse-quents)
14. rules = PruneRules(rules);
15. if (rules.size>minRules) then
16. return rules;
17. end if
18. support = support - delta;
19. end while

3.2 Experimental Evaluation

3.2.1 Evaluation Metrics

We assess the classifier in view of error rate with different prediction plans. We additionally report the accuracy rate. The error rate signifies the quantity of wrong predictions over the aggregate number of predictions. The accuracy rate means the quantity of right predictions over the aggregate number of predictions.

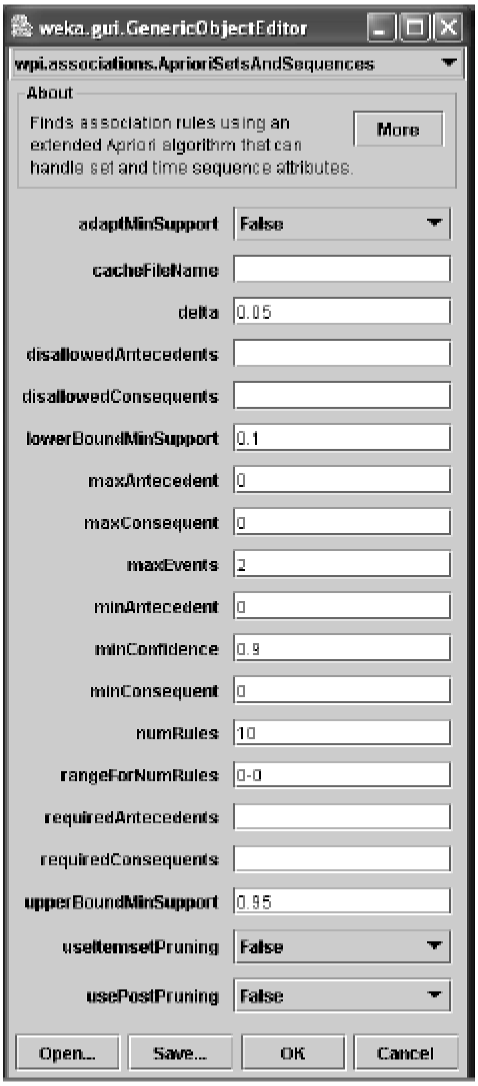

Figure 3.1: Parameter Menu for Our Extended Association Rule Mining

3.2.2 Experimental Results

We divide this section into two parts. In part 1, we focus on the improvements made to AprioriSets And Sequences by itemset pruning in the presence of constraints. Here we evaluate performance based on time taken for mining and generating rules and the number of maximal frequent itemsets generated. A frequent itemset is considered maximally frequent if none of its supersets is frequent.

Data Set

We tested the classification system with the following datasets obtained from the UCI Machine Learning Repository: census-income, mushroom and forest cover. Table 3.5 shows the properties of these datasets. As part of pre-processing, continuous valued attributes were discretized using WEKA's instance based discretization ltem with the number of bins set to 10.

Dataset	# attr	class	# class values	# instances
sonar	61	rocks/mines	2	208
census-income	15	income-level	2	32,561
mushroom	23	edible/poisonous	2	8,124
forest cover	17	forest cover type	7	74,056

Table 3.1: Dataset Properties Item set Pruning in the presence of Constraints

As a major aspect of our experiments, we were interested in contrasting thing set pruning versus non-pruning. We ran experiments with the mushroom, evaluation pay and backwoods cover data sets. We produced single and multiple constraint classification rules. We watched the subsequent parameters, for example, the quantity of thing sets delivered, number of maximal thing sets created and time taken for generating rules.

Table 3.1 demonstrates the parameters utilized as a part of running the experiments. In these experiments, the goal was to create whatever number standards as could be allowed with the help more prominent than or equivalent to 1%. The base confidence was set to half.

Prune	Req. Ant	Req. Con	item sets	Rules	Max. item sets	Time(s)
No	none	Class	45391	21101	158	4951
Yes	none	Class	42620	21101	42	4357
No	odor	Class	45391	8288	158	1153
Yes	odor	Class	33160	8288	26	1813

Table 3.2: Experimental Parameters

Prune	Req. Ant	Req.Con	Itemsets	Rules	Max. itemsets	Time(s)
No	None	Class	45391	21101	158	4951
Yes	None	Class	42620	21101	42	4357
No	Odor	Class	45391	8288	158	1153
Yes	Odor	Class	33160	8288	26	1813

Table 3.3: Comparison of Constraint-based Pruning vs. Non-Pruning for Mushroom Dataset

In Table 3.3, we show the results for the mushroom data set. The rest column appears if constraint based pruning was chosen or not. On account of pruning being exchanged o , all competitor thing sets are utilized as a part of generating valid thing sets at each level of the Apriori procedure. In looking at the rest two columns (single constraint), we watch the lessening in the quantity of thing sets delivered and the decrease in time taken for generating the guidelines. Be that as it may, interestingly, in the following two columns (twofold constraint) despite the fact that the quantity of thing sets delivered diminishes, the time taken increments. We gured this is where countless subsets are dropped from consideration because of the constraint based pruning. The nal output of the database for help of those things costs a huge time, expanding the general time.

Prune	Req. Ant	Req. Con	itemsets	Rules	Max. itemsets	Time(s)
No	none	class	1071	350	82	85
Yes	none	class	410	350	36	88
No	relationship	class	1071	22	82	87
Yes	relationship	class	100	22	31	16

Table 3.4: Comparison of Constraint-based Pruning vs. Non-Pruning for Census-Income Dataset

As seen in Table 3.3, in the case of a single constraint, the results for pruning and non-pruning are very similar. In the case of two constraints, the pruning leads to better performance in terms of time, approximately 1/5 of the time taken without pruning.

Prune	Req. Ant	Req. Con	itemsets	Rules	Max. itemsets	Time(s)
No	none	class	4297	1247	45	651
Yes	none	class	2673	1247	19	613
No	aspect	class	4297	144	45	678
Yes	aspect	class	605	144	22	208

Table 3.5: Comparison of Constraint-based Pruning vs. Non-Pruning for Forest-Cover Dataset

IV CONCLUSIONS

We developed a characterization of those thing sets that will conceivably frame rules that fulfill the given constraints. This characterization enables us to lter out from consideration all the item sets with the end goal that neither they nor any of their supersets will shape valid rules.

We developed a classification framework that is based on association rule mining in the WEKA condition. We implemented the CBA model building algorithm and compared the execution of CBA with All Rules Model (ARM) where all the mined rules are a piece of the model. We developed various modes to predict an unclassified case, for example, single rule or different rule prediction weighed by confidence/support.

REFERENCES

[1] RakeshAgrawal, TomaszImirlinksi, and ArunSwami.Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACMSIGMOD International Conference on Management of Data, pages 207-216, 1993.

[2] Sergio A. Alvarez. Binary search strategy for adaptive minimal support Personal Communication, 2002.

[3] Erik B. Sudderth, Antonio Torralba, William T. Freeman, and Alan S. Willsky. Describing visual scenes using transformed objects and parts.International Journal of ComputerVision, 77(1–3):291–330, 2008.

[4] JosefSivic and Andrew Zisserman. Video data mining using configurations of viewpoint invariant regions. In CVPR, pages 488–495, 2004.

[5] A. Thabtah. A review of associative classification mining. Knowledge Eng. Review, 22(1):37–65, 2007.

[6] RisivardhanThonangi and VikramPudi. ACME: An associative classifier based on maximum entropy principle. In ALT, pages 122–134, 2005.

[7] H. Verlinde, Martine De Cock, and R. Boute. Fuzzy versus quantitative association rules:A fair data-driven comparison. IEEE Transactions on Systems, Man, and Cybernetics–PartB: Cybernetics, 36(3):679–683, 2005.

[8] Eibe Frank and Ian H. Witten. Data Mining.Morgan Kaufmann Publishers, 2000.

[9] PradeepShenoy, Jayant R. Haritsa, S. Sudarshan, Gaurav Bhalotia, MayankBawa, and Devavrat Shah. Turbo-charging vertical mining of large databases. In SIGMOD Conference, pages 22–33, 2000.

[10] Ashok Savasere, Edward Omiecinski, and Shamkant B. Navathe. An efficient algorithm for mining association rules in large databases. In VLDB, pages 432–444, 1995.

[11] Eibe Frank and Ian H. Witten. Data Mining.Morgan Kaufmann Publishers, 2000.

[12] PradeepShenoy, Jayant R. Haritsa, S. Sudarshan, Gaurav Bhalotia, MayankBawa, and Devavrat Shah. Turbo-charging vertical mining of large databases. In SIGMOD Conference, pages 22–33, 2000.

[13] AshokSavasere, Edward Omiecinski, and Shamkant B. Navathe. An efficient algorithm for mining association rules in large databases. In VLDB, pages 432–444, 1995.

[14] William W. Cohen. Learning trees and rules with set-valued features. In Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Articial Intelligence Conference, pages 709{716, Menlo Park, August 4{8 1996. AAAI Press / MIT Press.

[15] MichelangeloCeci, Annalisa Appice, and DonatoMalerba. Spatial associative classification at di erent levels of granularity: A probabilistic approach. In PKDD, pages 99{111, 2004.

[16] Proc. 20th Int. Conf. Very Large Data Bases, VLDB, pages 487{499. Morgan Kaufmann, 12{15 September 1994.

[17] HerveJegou, MatthijsDouze, and CordeliaSchmid.On the burstiness of visual elements.In CVPR, pages 1169–1176, 2009.

[18] Fred´eric´ Jurie and Bill Triggs. Creating efficient codebooks for visual recognition. In ICCV,pages 604–610, 2005.

[19] Mehmet Kaya and RedaAlhajj. Integrating multi-objective genetic algorithms into clustering for fuzzy association rules mining. In Proceedings of the Fourth IEEE InternationalConference on Data Mining, ICDM, pages 431–434, 2004.

[20] Mehmet Kaya, RedaAlhajj, FarukPolat, and AhmetArslan. Efficient automated miningof fuzzy association rules. In DEXA, pages 133–142, 2002.

[21] Yan Ke and Rahul Sukthankar.Pca-sift: A more distinctive representation for local imagedescriptors. In CVPR, pages 506–513, 2004.

[22] Jingen Liu, Yusuf Aytar, Omer Bilal Orhan, Jenny Han, and Mubarak Shah.University of central florida at trecvid 2007 semantic video classification and automatic search. InTRECVID, 2007.

[23] Bing Liu, Wynne Hsu, and Yiming Ma.Integrating classification and association rulemining. In KDD, pages 80 – 86, 1998.

[24] Wenmin Li, Jiawei Han, and Jian Pei. CMAR: Accurate and efficient classification basedon multiple class-association rules. In ICDM, pages 369–376, 2001.

[25] David G. Lowe. Distinctive image features from scale-invariant keypoints. InternationalJournal of Computer Vision, 60(2):91–110, 2004.

[26] Svetlana Lazebnik, CordeliaSchmid, and Jean Ponce.Beyond bags of features: Spatialpyramid matching for recognizing natural scene categories.In CVPR, pages 2169–2178, 2006.