

A review on comparative performance analysis of associative classifiers

Zulfiqar Ali ^{1,2,*}, Waseem Shahzad ¹, Syed Khuram Shahzad ²

¹Department of Computer Science, National University of Computer & Emerging Science, Islamabad, Pakistan

²Department of Computer Science and Information Technology, The University of Lahore, Lahore, Pakistan

ARTICLE INFO

Article history:

Received 26 December 2016

Received in revised form

10 March 2017

Accepted 15 April 2017

Keywords:

Associative classification

Association rule mining

KEEL

Data mining

ABSTRACT

In this study we provided comparative study of associative classifiers which can be exploited for the discovery of business rules from the huge structured and unstructured data that can be used in the business analytic. Associative classification is a hybrid approach combining the classification rules mining and association rules mining that are two important data mining tasks. There are various emerging classification problems in various domains of knowledge like medical data, images, audio, video and textual data. Associative Classification approaches are exploited in various fields for the classification purposes. We compare the selective associative classification methods namely CBA, CBA2, CMAR-C, CFAR-C, CPAR-C, and Fuzzy-FARCHD-C by exploiting the implementation of these methods in KEEL data mining tool on public datasets. Our experimental results reveals that the performance of the Fuzzy-FARCHD-C is promising than other methods in terms of accuracy. The performance of the associative classifiers drastically decreases on the datasets with higher number classes and attributes.

© 2017 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The performance of any classification approach mainly depends on the accuracy and efficiency of the classifier. There are various emerging classification problems in various domains of knowledge like medical data, images, audio, video and textual data. The growth in the data reservoirs in the fields of business, science, stock exchange, basket analysis and geology is very high due to the availability of inexpensive storage resources. For example the sensors and data management systems enabled the medical researchers to gather voluminous data. The high growth ratio and huge data creates a challenging problem i.e. knowledge discovery from the huge databases in the field of Data Mining. For the appropriate, effective and comprehensive knowledge discovery for the managers and decision makers; researchers are proposing continuously more efficient knowledge mining approaches. The classification approaches have been studied mainly in the field of neural networks, expert systems, machine learning and statistics. There are various approaches exploited for the building of associative

classifiers. Ant Colony Optimization based Associative Classification approaches is prominent and promising in terms of accuracy and rules discovery. Here is some representative classification rule discovering approaches exploiting ACO for the Associative Classification. Ant Colony Optimization is prominently used for the discovery of classification rules and association rules which results in efficient, robust and more accurate classifiers. The ACO was first applied for the discovery of classification rules by [Parpinelli et al. \(2002\)](#) which is known as AntMiner. [Liu et al. \(2002\)](#) proposed extension in the basic AntMiner algorithm in AntMiener2 ([Liu et al., 2002](#)) an AntMiner3 in [Liu et al. \(2003\)](#). [Martens et al. \(2007\)](#) proposed an AntMiner+ algorithm based on Max-Min Ant System that differs from the previously proposed AntMiners in several aspects. [Shahzad and Baig, 2010](#) proposed improvements in the cAntMiner algorithm, that provided promising classification rule discovery in medical data sets. They proposed a new bio inspired hybrid classification approach, named ACO-AC in [Shahzad and Baig \(2011\)](#). ACO-AC algorithm exploits hybrid approach by combining the idea of association rules mining and supervised classification. [Baig and Shahzad \(2012\)](#) proposed another bio inspired classification approach, named AntMiner-C in [Baig and Shahzad \(2012\)](#). The literature study shows the application of ACO for the discovery of rules for the classification task using supervised training data.

* Corresponding Author.

Email Address: zulfiqar.ali@nu.edu.pk (Z. Ali)

<https://doi.org/10.21833/ijaas.2017.06.014>

2313-626X/© 2017 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Otero et al. (2008) proposed a classification rule mining ACO based algorithm which introduced improvements in Ant-Miner for coping with continuous attributes, named cAnt-Mine (Otero et al., 2008). Jin et al. (2006) proposed a new classification rule mining algorithm named ACO-Miner in Jin et al. (2006).

Thabtah (2007) provided the review of associative classification. Vyas et al. (2008) described the application of associative classifiers for Predictive analytics. Soni and Vyas (2010) surveyed the application of associative classifiers for predictive analysis in health care Data Mining. This study used the KEEL Data Mining Tool with more number of associative classifiers. In this article we investigate the performance of the Associative Classification (AC) methods namely CBA, CBA2, CMAR-C, CFAR-C, CPAR-C, and Fuzzy-FARCHD-C by exploiting the implementation of these methods in KEEL data mining tool by using the public datasets. Our experimental results show that the performance of the Fuzzy-FARCHD-C is better than the other methods in terms of accuracy.

The section 2 of the paper discusses the associative classification and describes the selective associative classification methods that are under the focus of this study for the comparative analysis. The section 3 explains the experimental Set-up exploited for this study, data sets and KEEL tool used for the experimentation. The Section 4 describes the comparative performance results and finally the last section 5 concludes the study.

2. Associative classification

Associative Classification is a classification approach which integrates the classification rules mining and association rules mining that are two important data mining tasks. The association rule mining is unsupervised learning in which no class attribute involved during the discovery of rules. The aim of the association rule mining is to discover associations between items in a transaction database. In association rule mining there could be more than one attribute in the consequent of a rule. The associative classification is a supervised learning where a class must be given for the discovery of classification rules.

The objective of associative classification approach is to construct a classifier that can forecast the classes of test data objects. The only class attribute is in the consequent of a rule. The over fitting is a considerable issue in the associative classification rule discovery. The over view of the selective Associative Classification approaches for the performance analysis under this study is given in the following sections.

2.1. CBA

Ma and Liu (1998) proposed a new hybrid classification approach by integrating concept of Association Rule Mining and Classification Rule

Mining that is named Classification Based on Associations (CBA). In this associative classification approach the integration is done by focusing on discovery of a special subset of association rules that are known as class association rules (CARs). For the discovery of all class association rules that satisfy the minimum support and minimum confidence constraints an existing association rule mining algorithm (Agrawal and Srikant, 1994) is exploited in this approach. The CBA associative classifier consists of two parts; a rule generator (CBA-RG) which is based on the Apriori algorithm and a classifier builder (CBA-CB). This approach possesses various advantages like the discretization of continuous attributes based on the classification pre-determined class target. The Data Mining task in CBA consists on the three steps; 1) discretization of continuous attributes, if any ; 2) generating all the class association rules; 3) building a classifier based on the generated class association rules.

2.2. CBA2

Liu et al. (2001) proposed the enhancement and improvements in an associative classifier CBA based on Ma and Liu (1998). The new improved associative classification approach is named CBA2 developed in (Liu et al., 2001). In this paper the authors tried to coup up with weaknesses of an exhaustive search based classification system CBA. The authors proposed two new techniques to deal with the observed weaknesses of the classification approaches. The first weakness observed is that as the traditional association rule mining exploits only a single minsup in rule generation which results inadequate for unbalanced class distribution. Secondly classification data often contains a huge number of rules, which may cause combinatorial explosion. For various databases, the rule generator is unable to generate rules with many conditions while such rules may be important for accurate classification. The first problem in this approach is tackled by using multiple class minsups in rule generation instead of single minsup as in CBA. The second problem which is caused by exponential growth of the number of rules is dealt indirectly. The decision tree method (Salzberg, 1994) is exploited.

The main working concept of the CBA2 is to use the rules of CBA2 to segment the training data and then select the classifier. These improvements in CBA improved the accuracy and lower error rate of the classification.

2.3. CMAR

Li et al. (2001) proposed a new associative classification method known as Classification based on Multiple Association Rules (CMAR). The CMAR associative classification approach is based on the frequent pattern mining method. This method extends FP-growth, constructs a class distribution-associated FP-tree. The CMAR applies a CR-tree structure to store and retrieve mined association

rules efficiently, and prunes rules effectively based on confidence, correlation and database coverage. The CMAR classification approach determines the class label by a set of rules instead of relying on a single rule for classification. For the improvement in accuracy and efficiency CMAR employs a novel data structure named CR-tree. The CR-tree is exploited in CMAR to compactly store and efficiently retrieve a large number of rules for classification. CMAR consists of two phases: rule generation and classification. The associative classification approach CMAR possesses capabilities to mine large database efficiently.

2.4. CPAR

Yin and Han (2005) proposed a new classification approach which combines the advantages of both associative classification and traditional rule-based classification namely known as Classification based on Predictive Association Rules (CPAR). The CPAR exploits a greedy algorithm for the generation of rules directly from the training data. CPAR inherits the basic idea of First Order Inductive Learner (FOIL) (Jin et al., 2006) in rule generation and integrates the features of associative classification in predictive rule analysis. The CPAR possesses distinguishing features with respected other associative classification approaches like; 1) generates a much smaller set of high-quality predictive rules directly from the database; 2) avoids to generating redundant rules; 3) it uses the best k rules for predicting the class label of an example. The repeated calculations are avoided by using dynamic programming in the CPAR approach.

2.5. Fuzzy-FARCHD

Alcala-Fdez et al. (2011) proposed a Fuzzy Association Rule-based Classification method for High-Dimensional problems (FARCHD). This approach targets the problem of exponential growth of the fuzzy rule space faced during the inductive learning of fuzzy rule based classification systems. The FARCHD approach promises to reduce problems of scalability and complexity of the classification process. The FARCHD classification approach consists of three stages. The first stage is fuzzy association rule extraction for classification. In this stage a search tree is employed to list all possible frequent fuzzy item sets. The second stage is the candidate rule prescreening. The candidate rule prescreening decreases the computational cost of the genetic post-processing stage. The third stage of FARCHD is Genetic rule selection and lateral tuning. This stage is exploited for the selection and tuning a compact set of fuzzy association rules with high classification accuracy of the classifier. This approach obtains accurate and compact fuzzy rules which results in a classifier with a low computational cost.

2.6. CFAR

Chen and Chen (2008) proposed an associative classification approach namely Classification with Fuzzy Association Rules (CFAR). The CFAR approach exploits fuzzy logic that is suitable to deal with the “sharp boundary” problem by providing a flexible and intelligent remedy. The classical association rules are special cases of fuzzy association rules. The semantics of a fuzzy association rule is richer and natural language nature which are more promising. The fuzzy associative rules based associative classifiers are more promising to mine larger datasets with quantitative domains and to generate classification rules with more general semantics and linguistic expressiveness. The Classification with Fuzzy Association Rules approach has better understandability in terms of the number of rules and the smooth boundaries with respected to other state-of-the art associative classifiers, while keeping the accuracy equally satisfactory.

3. Experimental set-up

In this section, we conduct experiments to evaluate the performances of the associative classification systems. For the comparative performance analysis of the selective associative classifiers we exploited the implementations of these algorithms included in KEEL (Alcalá et al., 2010). The overview of the Data Mining and machine learning tool KEEL is given in the following section. In this section we describe the datasets used for the comparative analysis of the associative classifier in terms of accuracy. The parameters set for the experiments and the experiment graph designed for these experiments in the KEEL tool are described in this section.

3.1. Data sets

The description of datasets used for the comparative performance analysis of the selective associative classifiers under this study is given in the Table 1. The number of attributes (#Attributes), number of instances in the database (#Examples) and number of classes (#Classes) are shown in the table. The missing values (Missing_V) in the dataset are representing by “Yes” (missing values present)/ “No” (missing values not present). The missing values of the datasets are imputed with the KMean-MV module implemented in KEEL. The datasets are discretized with the Ameva-D module included in KEEL as the associative classifiers accept the discretize form of datasets.

We use the 10-fold cross-validation model for the datasets provided in KEEL. Table 1 summarizes the main characteristics of the 12 datasets which are given at Knowledge Extraction based on Evolutionary Learning (KEEL)-dataset repository (Alcalá et al., 2010).

3.2. KEEL

The Knowledge Extraction based on Evolutionary Learning (KEEL) (Alcalá et al., 2010) is an open source software tool to assess Evolutionary Algorithms for data mining problems including regression, classification, clustering, pattern mining and so on. The screenshot of KEEL data mining tool version 3.0 is shown in Fig. 1. This tool provides a simple GUI based on data flow to design experiments with different datasets. KEEL provides a good collection of computational intelligence algorithms which can be used by the researchers in order to assess the behavior of the algorithms. Moreover it may also be used to compare new proposed techniques with the state-of-the art approaches of their corresponding areas.

3.3. Experiment graph

The experiment graph shows the components of the experiment and describes the relationships between them. The experimental graph of the comparative study is given in the Fig. 2. The first component of the experimental graph is data which enables to select the datasets given in the KEEL Tool as well as to load user datasets. In our study, we selected standard KEEL datasets. The second component of the graph is KMeans-MV which is a module to impute the missing values in the database.

The third component of the experiment graph is module for data discretization.

Table 1: Data sets considered for the experimental study

Dataset	#Attributes	#Examples	#Classes	Missing_V
Bupa	6	345	2	No
Cleveland	13	297	5	Yes
Ecoli	7	336	8	No
Glass	9	214	7	No
Haberman	3	306	2	No
Iris	4	150	3	No
Monks	6	432	2	No
New-Thyroid	5	215	3	No
Pima	8	768	2	No
Vehicle	18	846	4	No
Wine	13	178	3	No
Wisconsin	9	683	2	Yes



Fig. 1: The screenshot of KEEL data mining tool version 3.0

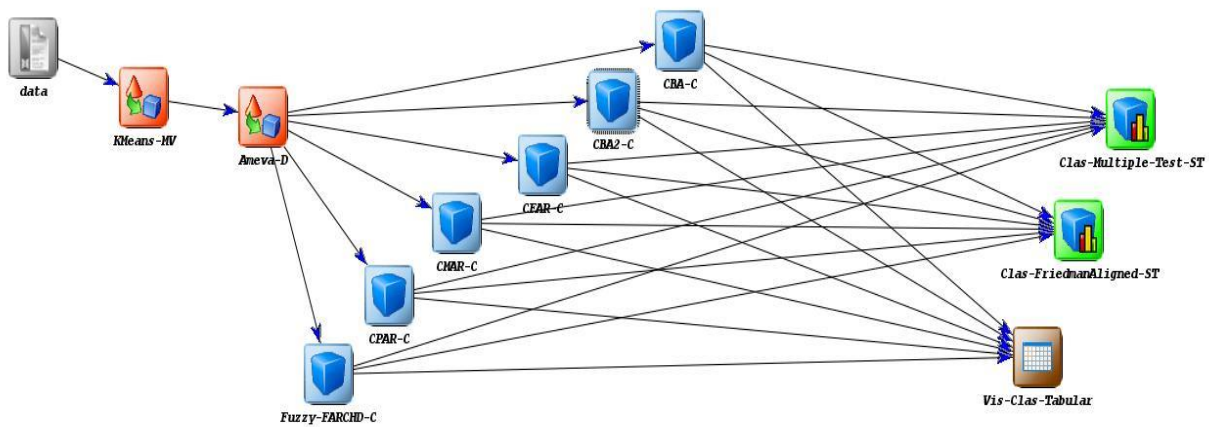


Fig. 2: The experiment graph

In our case we use the Ameva-D module for the discretization of continuous data values. The fourth stage of the experiment graph is Associative Classification methods which are under the focus of study. Here we selected six algorithms of the family of Associative Classification i.e., CBA, CBA2, CMAR-C, CFAR-C, CPAR-C and FARCD-C. The last stage of the experiment graph is the modules for the representation of the results of classifier and statistical module for the analysis of the results produced by the algorithms used in the experiment.

3.4. Parameters of the methods

The parameters of the associative classifiers under the focus of this comparative study are shown

in the Table 2. The parameters of the methods are selected according to the recommendation of the corresponding authors within each proposal which are the default parameters settings included in the KEEL software tool (Alcalá et al., 2010).

In the Table 2, *Minconf* stands for minimum support, *Minconf* for minimum confidence, and *RuleLimit* for maximum candidate rules limit in the corresponding methods. The cut threshold is represented by *MS* in method CFAR-C. The database coverage threshold (delta) is represented by δ in CMAR-C method. For the CPAR-C method weight decay factor (alfa) is represented by α and where *k* is the number of rules used in prediction. For FARCD-C *Depthmax* shows the depth of the trees; *k* shows the parameter of the prescreening; *Pop* for

population size and *BITSGENE* represents the bits per gen.

Table 2: Parameters of the methods for experiment

Methods	Parameters
CBA-C	<i>Minsup</i> = 0.01, <i>Minconf</i> = 0.5, <i>Pruned</i> = yes, <i>RuleLimit</i> = 80,000
CBA2-C	<i>Minsup</i> = 0.01, <i>Minconf</i> = 0.5, <i>Pruned</i> = yes, <i>RuleLimit</i> = 80,000
CFAR-C	<i>Minsup</i> = 0.1, <i>Minconf</i> = 0.85, <i>MS</i> = 0.15
CMAR-C	<i>Minsup</i> = 0.01, <i>Minconf</i> = 0.5, $\delta\delta = 4$,
CPAR-C	$\delta\delta = 0.05$, <i>min_gain</i> = 0.7, $\alpha = 0.66$, $k = 5$
FARCHD-C	<i>Minsup</i> = 0.05, <i>Maxconf</i> = 0.8, <i>Depthmax</i> = 3, $k_t = 2$, <i>Pop</i> = 50, <i>Evaluations</i> = 15,000, <i>BITSGENE</i> = 30, $\delta\alpha = 0.15$

4. Experimental results

Table 3 shows the comparative performance of the selected associative classifiers. We use the implementation of the corresponding algorithms in KEEL 3.0 for our comparative performance analysis of the associative classifiers family. The values bold

face shows the wining of the corresponding Associative Classifier. The performance of FARCHD-C is overall better than other associative classifiers under focus of this study. Fig. 3 shows the comparative analysis of the associative classifiers in terms of accuracy. There is very interesting pattern emerged in the results of these classifiers with the comparative analysis in terms of accuracy. All methods show the lower performance on Glass and Vehicle datasets comparatively to the other databases. The performance of CMAR-C and CFAR-C drastically decreases on the datasets Glass and Vehicle respectively in terms of accuracy. The performance of all associative classifiers is same on Cleveland, Iris, Monks, New-Thyroid and Wisconsin datasets. By the critical observations of the results it reveals that the performance of the associative classification methods used in this study decreases with the increase in number of attributes and number of classes in the databases.

Table 3: The comparative performance results of associative classification methods (accuracy in %)

Dataset	CBA-C	CBA2-C	CMAR-C	CFAR-C	CPAR-C	FARCHD-C
Bupa	66.9423	67.5053	68.0851	53.8756	65.1342	66.6734
Cleveland	66.9423	53.7048	56.4321	54.3500	54.9365	56.4712
Ecoli	66.9423	72.1925	73.5456	47.9096	71.1149	75.4416
Glass	66.9423	49.7675	20.3936	36.5606	45.6770	50.4431
Haberman	66.9423	74.1349	72.3656	45.7087	73.5484	72.6393
Iris	66.9423	92.7273	91.5152	92.7273	95.7576	93.9394
Monks	66.9423	97.5159	97.5159	97.5159	97.5159	97.5159
New-Thyroid	66.9423	94.0968	93.7033	93.6639	91.9717	94.5100
Pima	66.9423	72.1981	71.8408	65.0924	71.0221	73.6335
Vehicle	66.9423	70.6672	70.1350	40.8085	63.0481	70.5730
Wine	66.9423	92.3054	94.8604	92.3351	95.9299	94.4147
Wisconsin	66.9423	95.5656	95.9627	95.3134	94.1408	96.4860
Average	66.9423	77.6984	75.5296	67.9884	76.6497	78.5617

Table 4 shows the record of Win/Draw/Loss of the AC classifiers. The CBA wins 2 times and loss 10 time in term s of accuracy with other classifiers. The CBA2, CMAR-C and CPAR-C 1 time wins, 1 time draw and 10 losses in the comparison with other approaches. The CFAR-C draws 1 time but does not win even one time from the other classifier in terms of accuracy. The FARCHD-C performed better than other classifiers include in this study. The FARCHD-C win 5 times, draw 1 and loses 6 times as given in the Table 4.

4.1. Results and analysis of associative classifiers on various databases

This sections represents the performance variation of associative classifiers on the various databases under the discussion in graphs to increase the understandability. Fig. 3 shows the collective performance results of associative classifiers under focus of this study on various public datasets i.e., Bupa, Cleveland, Ecoli, Glass, Haberman, Iris, Monks, New-Thyroid, Pima, Vehicle, Wine, and Wisconsin.

With Bupa dataset, the performance of all the classifiers under focus is similar except CFAR-C. The performance behavior of ACs on Cleveland datasets is very interesting. There is significant variation in the performance of associative classifiers in terms of

accuracy. The CMAR-C and FARCHD-C provided more promising results as compared to others. On Ecoli database the classifier "CFAR-C" remained behind the other classifiers.

On Glass dataset the performance of CBA-C, CBA2-C and FARCHD-C remained same level while the performance of CMAR-C drastically lower with respect to other approaches. On Haberman database the performance remained same except CFAR-C. The performance of associative classifiers on Iris dataset significantly various and CPAR-C produced promising results.

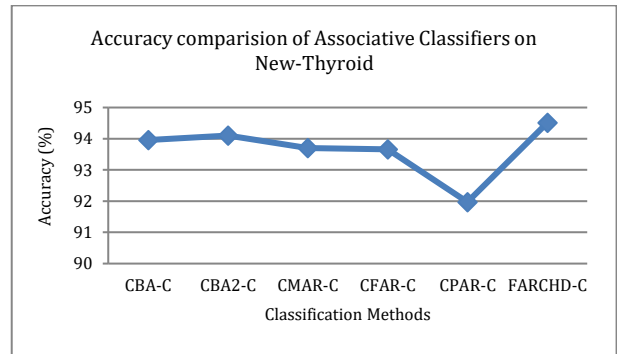
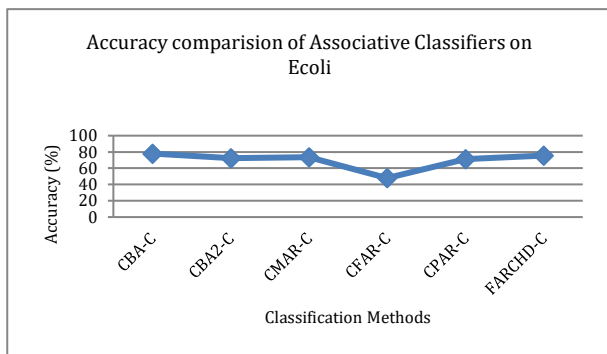
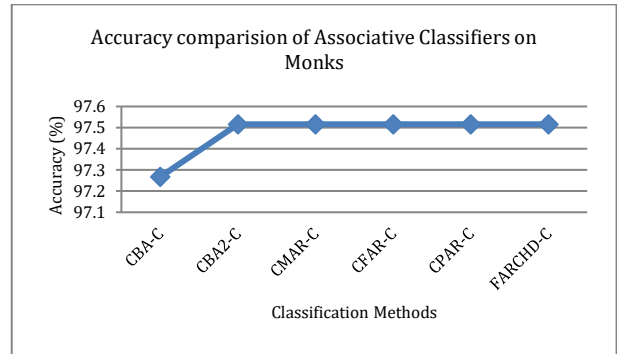
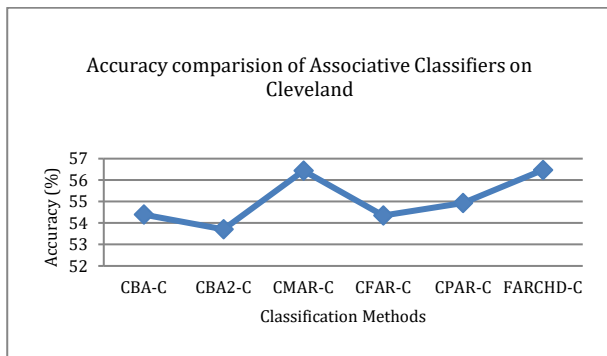
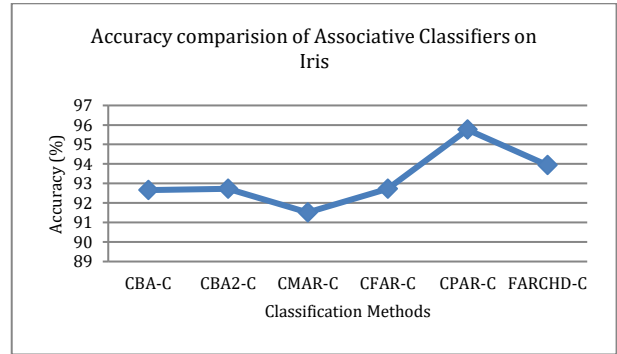
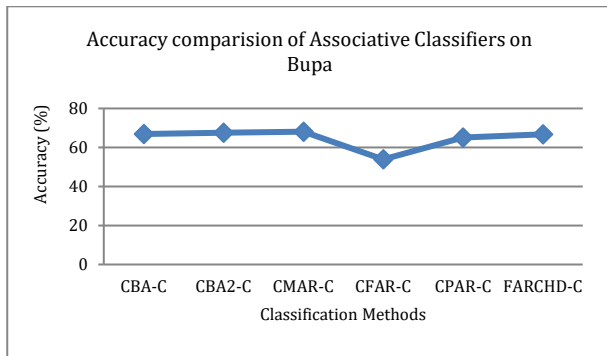
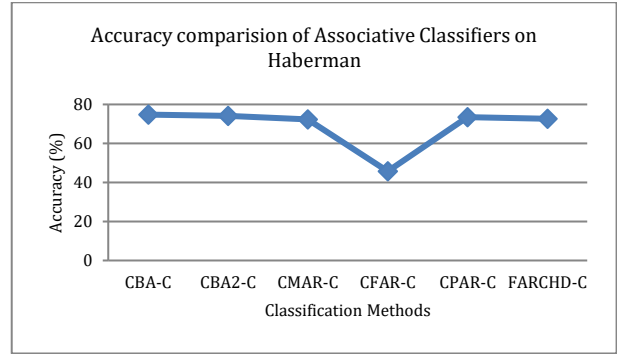
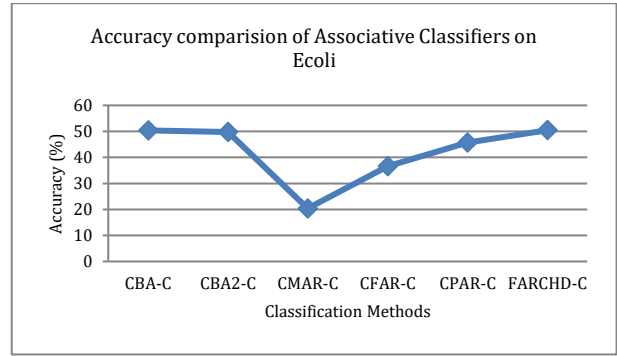
On Monks dataset the performance of all the classifiers remains same. The classification results in terms of accuracy for New-Thyroid dataset remained same for all the associative classifiers except CPAR-C. The performance of CPAR-C Associative Classifier is lower than others.

The accuracy results of associative classifiers for the Pima database reveals that the performance of FARCHD-C are more promising as compared to others while the performance of CFAR-C is lower than the other approaches as shown in the Fig. 3. The performance of CFAR-C significant decrease on Vehicle database with respect to other associative classification approaches that are under the focus of this comparative study.

The classification results of associative classifiers on Wine datasets are very interesting. The performance of CBA, CMAR-C and FARCHD-C are similar while CBA2 and CFAR-C remained behind the others. The highest performance is CPAR-C for the Wine database. The performance of CPAR-C is lower than the other approaches on the Wisconsin dataset. With the critical observation of the results it is revealed that there is variation in performance of associative classifiers on variation of the database.

Some associative classification approach performs well on one database while bad on the other database.

On the average from the Table 4, we conclude that FARCHD-C classifier performed better than the other approaches. The CBA classifier is at the second position in this comparative study in terms of accuracy on the public data sets. CBA is winner 2 times while CFAR-C produced lowest results with respect to other complete approaches as shown in the Table 4.



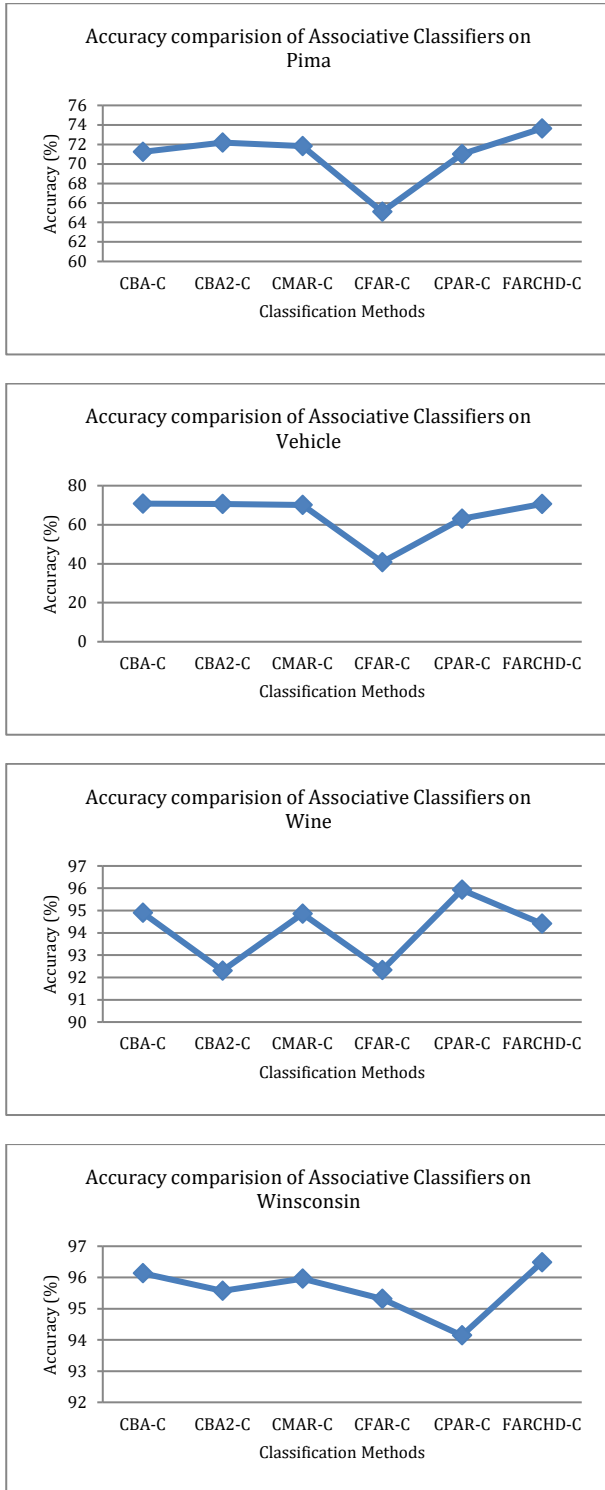


Fig. 3: Performance analysis of associative classifiers at various databases

Table 4: Win/Draw/Loss record of associative classifiers

	CBA-C	CBA2-C	CMAR-C	CFAR-C	CPAR-C	FARCHD-C
Win	2	1	1	0	1	5
Draw	0	1	1	1	1	1
Lose	10	10	10	11	10	6

5. Conclusion

This article focuses on the performance analysis of the Associative Classification approaches. Associative Classification is a hybrid approach

combining the classification rules mining and association rules mining that are two important data mining tasks. Associative classification approaches are exploited in various fields for the classification purposes. We compare the selective associative classification methods (CBA, CBA2, CMAR-C, CFAR-C, CPAR-C, and Fuzzy-FARCHD-C) by exploiting the implementation of KEEL data mining tool by using the public datasets. The performance of the Fuzzy-FARCHD-C is promising than other methods in terms of accuracy. The CMAR-C significantly degrades on the Glass datasets while the performance of CFAR-C decreased on the Vehicle dataset. The performance of the associative classifiers under study significantly degrades on the databases with increase in number of attributes and number classes.

In future we will analyses the performance of associative classifiers considering other parameters and derive the significance of results by using statistical methods.

Acknowledgement

This research work is sponsored by the Higher Education Commission of Pakistan in the form of Scholarship for the Ph.D. program.

References

Agrawal R and Srikant R (1994). Fast algorithms for mining association rules. In the 20th International Conference on Very Large Data Bases, VLDB, 1215: 487-499.

Alcalá J, Fernández A, Luengo J, Derrac J, García S, Sánchez L, and Herrera F (2010). Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. Journal of Multiple-Valued Logic and Soft Computing, 17(2-3): 255-287.

Alcala-Fdez J, Alcalá R, and Herrera F (2011). A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning. IEEE Transactions on Fuzzy Systems, 19(5): 857-872.

Baig AR and Shahzad W (2012). A correlation-based ant miner for classification rule discovery. Neural Computing and Applications, 21(2): 219-235.

Chen Z and Chen G (2008). Building an associative classifier based on fuzzy association rules. International Journal of Computational Intelligence Systems, 1(3): 262-273.

Jin P, Zhu Y, Hu K, and Li S (2006). Classification rule mining based on ant colony optimization algorithm. In: Huang D, Li K, Irwin GW (Eds.), Intelligent Control and Automation: 654-663. Springer, Kunming, China.

Li W, Han J, and Pei J (2001). CMAR: Accurate and efficient classification based on multiple class-association rules. In the IEEE International Conference on Data Mining (ICDM 2001), IEEE, San Jose, USA: 369-376. <https://doi.org/10.1109/ICDM.2001.989541>

Liu B, Abbass HA, and McKay B (2002). Density-based heuristic for rule discovery with ant-miner. In The 6th Australia-Japan Joint Workshop On Intelligent And Evolutionary System, School Of Computer Science, Australian Defence Force Academy, University of New South Wales, Australia.

Liu B, Abbass HA, and McKay B (2003). Classification rule discovery with ant colony optimization. In the IEEE/WIC International Conference on Intelligent Agent Technology, IEEE, Halifax Regional Municipality, Canada. <https://doi.org/10.1109/IAT.2003.1241052>

- Liu B, Ma Y, and Wong CK (2001). Classification using association rules: weaknesses and enhancements. In: Grossman R, Namburu R, Kamath C, Kegelmeye P (Eds.), *Data Mining for Scientific and Engineering Applications*: 591-605. Springer Science and Business Media, Berlin, Germany.
- Ma BLWHY and Liu B (1998). Integrating classification and association rule mining. In the 4th International Conference on Knowledge Discovery and Data Mining, AAAI Press: 80-86.
- Martens D, De Backer M, Haesen R, Vanthienen J, Snoeck M, and Baesens B (2007). Classification with ant colony optimization. *IEEE Transactions on Evolutionary Computation*, 11(5): 651-665.
- Otero FE, Freitas AA, and Johnson CG (2008). cAnt-Miner: an ant colony classification algorithm to cope with continuous attributes. In the International Conference on Ant Colony Optimization and Swarm Intelligence, Springer Berlin Heidelberg, Brussels, Belgium: 48-59. https://doi.org/10.1007/978-3-540-87527-7_5
- Parpinelli RS, Lopes HS, and Freitas A (2002). Data mining with an ant colony optimization algorithm. *IEEE Transactions on Evolutionary Computation*, 6(4): 321-332.
- Salzberg SL (1994). C4. 5: Programs for machine learning by j. ross quinlan. Morgan Kaufmann Publishers, inc., 1993. *Machine Learning*, 16(3): 235-240.
- Shahzad W and Baig A (2011). Hybrid associative classification algorithm using ant colony optimization. *International Journal of Innovative Computing, Information and Control*, 7(12): 6815-6826.
- Shahzad W and Baig AR (2010). Compatibility as a heuristic for construction of rules by artificial ants. *Journal of Circuits, Systems, and Computers*, 19(01): 297-306.
- Soni S and Vyas O (2010). Using associative classifiers for predictive analysis in health care data mining. *International Journal of Computer Applications*, 4(5): 33-37.
- Thabtah F (2007). A review of associative classification mining. *The Knowledge Engineering Review*, 22(01): 37-65.
- Vyas R, Sharma LK, Vyas OP, and Scheider S (2008). Associative classifiers for predictive analytics: Comparative performance study. In the 2nd UKSIM European Conference on Computer Modeling and Simulation, IEEE: 289-294. Liverpool, UK. <https://doi.org/10.1109/EMS.2008.29>
- Yin and Han (2005). Efficient classification from multiple heterogeneous databases. In the Ninth European Conference on Principles and Practice of Knowledge Discovery in Databases. pp. 404-416.