ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

MULTI ATTRIBUTE SIMILARITY INDEX DATA PRESENTATION FOR UNCERTAIN CATEGORICAL DATA

D. Veeraiah
Department of CSE, Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh, India
E-Mail: veeraiahdvc@gmail.com

ABSTRACT

Data summarization in unrealistic or uncertain data streams is a basic concept in relational data sources. For outstanding data summarization on uncertain data stream evaluation with jumps of data streams environments. Traditionally single attribute summarization approach was introduced to define related instances to construct Uncertain One Class Classifier to summarize class instances perfectively. This framework kernel density based method to generate possible score to obtain each attribute with feasible data maintenance; UOCC also provides support vector machine (SVM) representation to summarization concept based on user's preferences and user's requirement in stored data source. It was generated possible score based on data instances. It is failed to support data exploration based on data attributes (characteristics) to utilize data instances with cluster relational data sets. So, we propose to develop Multi Attribute Grouping Method (MAGM) to define data summarization and portioned attribute selection for data exploration in uncertain data streams. MAGM defines a matrix to construct unidentified records into cluster in uncertain reliable data streams with attribute partitioning and feature selection. Our experimental results show effective data summarization with uniform user's data exploration with their search histories from uncertain data streams with respect to time and other feature factors.

Keywords: K-Means, uncertain one class classifier, multi attribute, support vector machine, feature representation.

1. INTRODUCTION

Data mining is an aggressively concept in information retrieval based on different attributes from different data sources. For effective data collection from data sources with respect to relevant data one class learning is required to perform labelled based classification with individual training sequences on attributes. For some real world data outsourcing real time data set portioning with abnormal behavioural class label instances with expensive impossible data presentation. To learn these types of collective sequences in real time data set proceedings to classify target data into distinct classifier data procedures. For variety of different applications anomaly detection, document classification image annotation and content specification for different data formations.

By seeing above discussion, we find the issue of single attribute on vague subtle elements sources and thought synopsis considering of the client from record points of interest sources. Ordinarily prescribe a structure, known as vague one-class contemplating and thought synopsis considering structure (UOLCS) on misty points of interest sources, which manages subtle elements of uncertainty and the thought rundown examining in hazy one-class subtle elements sources. UOLCS includes two sections. In the primary angle, assemble a One Class Classifier for Uncertain (UOCC) data to the independent points of interest into the single attribute SVM contemplating stage to manufacture a superior classifier. In the second viewpoint, we audit client's thought move from points of interest sources by making a support vectors (SVs) - based bunching procedure over the record segments. To increase multi class label presentation with high dimensional data in real time data applications, a better system is required to process different attributes. So in this paper, we prescribe to create Multi Attribute Grouping Methods (MAGM) to characterize record joins in light of properties in indeterminate information streams with possible and ID formal parameters.

Thus, the effectiveness of current gathering accumulation methods may subsequently be disintegrated the same number of framework records are left unidentified. This paper introduces Multi Attribute Grouping Method (MAGM) procedure to enhance irregular lattice to give extensively low level data set representation. It is connection based approach to access irrelevant data present in grouped data with different attributes based on similarity features. This exploration only associates the hole between the procedure of data bunching and that of web connection investigate. It additionally expands the capacity to accumulation system for specific data, which has not acquired much consideration in the artistic works. Strategy of the bunch gathering approach appeared in Figure-1 with relative components in group social information bases.

VOL. 13, NO. 9, MAY2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

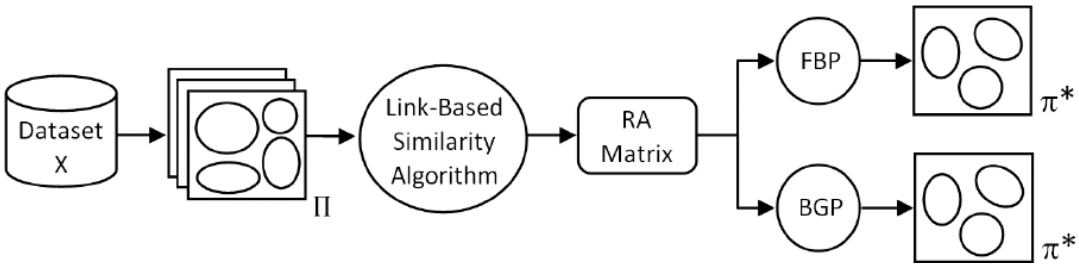

Figure-1. Procedure of the cluster ensemble for feature links in uncertain cluster data.

Notwithstanding the issue of grouping specific data that is analysed thus, the proposed structure is by and large with the end goal that it can likewise be effectively used to other data sorts.

The fundamental commitment of our proposed approach as takes after:

a) The component based procedure that changes over the issue of gathering outfits to bunching absolute information (i.e., aggregate marks)

b) The quick procedure that finds a definitive segment through relabelling the base clustering comes about

c) Edge and Graph-based techniques that utilization a diagram apportioning strategy

d) The sets insightful similitude methodology that uses co-event communication between data focuses.

Organization of this paper presents related work to define one class classification procedure to define attributes based on instances in section 2. Section 3 formalizes problem definition in one class classification in uncertain data streams. Section 4 defines cluster ensemble approach to define relations between selected features in cluster relational databases. Section 5explains experimental evaluation with respect to UOCC and CEA based on selected features. Sections 6 concludes overall conclusion about CEA approach to construct data summarization based on user group based on instances.

2. BACKGROUND WORK

This section describes different author's opinion in data exploration with different attributes with different data instances from different data sources. Some of research professors and authors explore their definitions regarding data retrieval from various data sources.

Aggarwal uses approach based approach to manage handle screw up slanted and missing data. The system focuses the issue of packing vague articles whose ranges are delineated by probability thickness limits and uses Voranoi describes attribute relations and R-Tree data to manage questions in relational data. For efficient attribute selection and classification with Support Vector Machine to questionable data present in randomly generated for two and single attributes on centre of interest points with confident relations. Like this Geo and Wang process query able independent and mistake able and undefined querying data. Tang *et al* defines display approaches to collect efficient querying in relational data. Based on investigation present in [13] to randomly generate fixed data attributes with repeated attributes in real time attribute partition. Like in [9] formalize particular alliance from different random possible entity semantic relations. For efficient mining of different attributes based on procedures with real time examples. Macular et.al defines combined data from randomly generated data relations with spatial database relations.

For automatic machine learning procedures with query able for undefined information relations inspected data streams gathering through large amount of information. For efficient attribute collection traditionally use UOCC with machine learning to adjusted results, first collect nearest attributes based on score with different similarity in uncertain single attribute for each relation, secondly progress quadratic programming in different relations based on SVM classification with summation based refinement for each query based on single query presented for each user present in relational database. To support different attributes with different relations in single class attributes shown in Figure-2.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

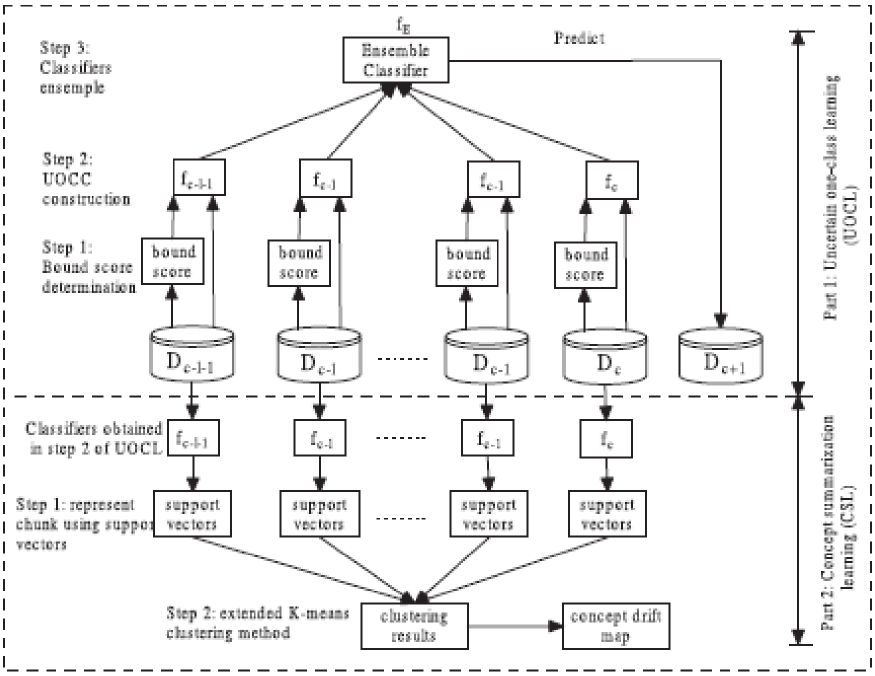

Figure-2. Concept summarization and one class learning in cluster data sets.

UOLCS structure comprises of two sections, the initial segment is to develop dubiously single node classifier from unverifiable information streams, and the remaining is idea outline machine sequence learning over the previous data streams. Two modules used in this scenario, they are 1) Single Attribute Learning 2) Topic Based Data Summarization.

2.1 Single attribute learning: One class learning approach defines three main modules in developing application for uncertain data streams with feasible data streams. For generate threshold score for instance based with local behavior using local attributes formed in different streams. In final step, for threshold based score generation to identify uncertain single attribute classification with repeated in undefined information streams. After classification different features related to information based on dimensionality in single node classification to extract relevant data streams in different data stream evaluations.

2.2. Topic based data summarization: In single attribute machine classification for learning, it is important to find different ideas with their independent common properties client from history pieces. Traditionally, advance our bolster similarity based grouping strategy for idea synopsis gaining to different relations. Normally, for efficient data relations present in general data processing based on similarity index with different utilization, to exploit client data relations with respect to similar features in relevant data assessments. To exploit efficient time processing with different relations for continues excess of streaming with verifiable information. Another approach utilizes highlight based grouping method to condense idea of the client. It first extricates highlights from an information lump and considers this piece as a virtual specimen spoke to by the separated components, hence; the entire information is introduced for specimen set, in that each virtual example speaks to one information piece.

These two steps are used to define one class classification procedures for threshold score calculation and define summarization based on classification with processing instances. This procedure achieves one class classification based on instances only. So a better system is required for classify with preferable summarization attributes with characteristics with reliable uncertain data streams. So next section defines those relations with realistic summarization from real data sets.

3. MULTI ATTRIBUTE GROUPING METHOD

In this section, we discuss about multi-attribute clustering specification, this approach actual listening with different attributes.

3.1. Basic procedure for data summarization

Let $C = \left(c1; c2; \ . \ . \ . \ ; \ cN\right)$ be a combination of data relations with N details factors and $\gamma = (\gamma 1, \gamma 2, ..., \gamma n)$ Ng be a team selection with M cluster analysis, each of which is referred to as a selection individual. Each platform clustering earnings a combined with categories. $\pi_i = \{X_1^i, X_2^i, X_3^i, ...X_n^i\}$, such that $\bigcup_{j=1}^{k_i} C_j^i = C$, where ki is different selection of cluster with different parameters. For each x in relational factor 2C with different characteristics characterizes the combined

ARPN Journal of Engineering and Applied Sciences

brand similarity with factor c with cluster sequence. In the i[th] similar grouping $X(x) = "j"(or"X_j^i")ifc \in X_j^i$. This partition gives primary assets π* of a complete set C, which contains grouped attributes with same attributes π [6][1].

So the basic cluster formation from different attribute clusters with suitable data with consensus learning functions based on results with similar attributes procedure shown in Figure-3.

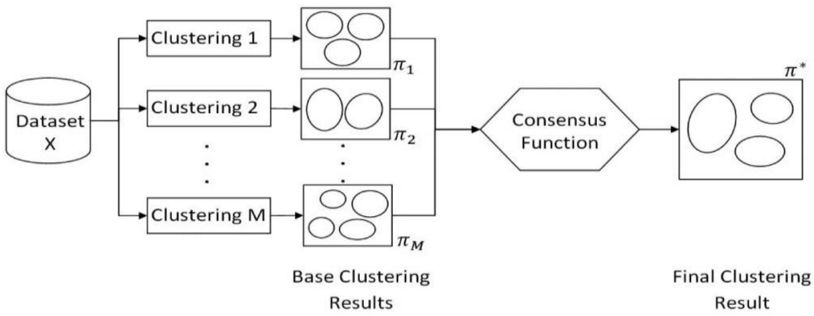

Figure-3. Cluster formation with different attributes with different Cluster in similar attribute partitioning.

3.2. Grouped creation approach: It is the basic concept to form different attributes in combination with same relations. In clustering, individual attributes over additional data streams. Selected attributes believe many conditions with similar features based on client requirements. In this situation, selected clients operate the overall system improvement based on cluster results. Consequentially, several attributes recommended present attributes in grouping approaches with range of particular successive relational attributes. Finally successive features were used to describe particular grouping requirements with different multi objectives.

3.3.Consensus functions: Out of overall attributes, randomly select grouped features have been designed for available information with attribute partition. Using markov chain matrix formation have similar attributes arranged in cognitive functions. Some of the feature based approaches with cluster analysis transforms operating attributes in real time data streams for detailed categorization. In Conesus, matrix formation with direct and indirect labelled formations.

3.4. Direct technique: In direct approach, depending attributes are individuals for selecting relational label i.e. π and number of attributes in π_I with multi objectives in relations for different formations using consensus function $\pi 1, \pi 2, \pi 3, ...\pi m$. To provide similar attributes with grouping for random selection from different data sets. In Markova chain model matrix different formations with attributes based on Euclidian distance between all attributes in data streams [8][9].

3.5.Outlier data cluster for attributes: From the procedure of direct technique with matrix formation and attribute arrangement with similar attributes in relations. Outlier formation based on attributes with multiple objectives in different consensus for grouping selected features in recent attributes to detect outlier from relations.

3.6. Algorithm for classification instances: For classification instances on attributes in real world synthetic data sets. Procedure of the classification instances with different attributes is as follows:

Table-1. Procedure of the proposed approach algorithm with relevance and label representation in training and testing usual representation.

Input: Set of Document s (D=d1,d2,d3,.......d_n), Set of queries (Q=q1,q2,q3,......q_n)
Output: Probability time (T=t1, t2, t3... t_n) stamp crime patterns results for submitted query (q).
Step 1: Initially query sequence (q_S=0).
Step 2: After getting query with time in documents the probability is p(q/t).
Step 3: Classify crime pattern sequences (C q_S) based on stored crime patterns.
Step 4: Calculate query frequency based on crime patterns with q_S.
Step 5: Query histogram with publication time (t) with increased documents (D)
Step 6: Compare each q_S probability time with original crime pattern time efficiency.
Step 7: Store probability of q_S in published crime documents
Step 8: Return each probability p(q/t) for each document (d) to query (q)

Multi Attribute Grouping Methods (MAGM) procedure shown in Table-1, it is step by step process for query attribute partition with multiple objects in relational data streams.

VOL. 13, NO. 9, MAY2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

4. EXPERIMENTAL EVALUATION

We formulate the effective evaluation analysis of proposed approach MAMG with comparison of UOCC on real data streams. To develop this application, we use JDK and Net beans for user interface construction to upload data sets and perform single attribute classification and multi attribute object classification from real time data streams. Sample data sets shown in Figure-4.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	LIVIO	Hemang	Male	########	livio_hem	208fa1780	91-829876	71	5	57	62	25	23	48
2	TYLAN	Apparajitc	Male	########	tylan_app	fbac38614	91-837531	89	5	18	23	8	1	9
3	Rajat	Vajramani	Male	6/5/2009	rajat_768(	4cc7ec9aa	91-838282	78	39	39	78	29	1	30
4	Madhulat:	Y	Female	########	madhulat:	12f5ad0d1	91-752958	2	77	93	170	59	4	63
5	Kanchan	CHARLY	Female	########	charly@gr	c4933ca23	NA	0	0	0	0	0	0	0
6	Sai	D	Female	########	d_236@gr	ba0028f0c	91-979586	19	2	4	6	73	86	159
7	Urmila	CORBEN	Female	########	urmila_co	bfa38b92c	91-984556	4	9	9	18	3	2	5
8	Shevanti	Q	Female	########	q@yahoo.	67feafb25	91-837287	98	4	17	21	8	58	66
9	Aruna	KELECHI	Female	########	aruna_kel	33c3649da	91-759243	2	71	1	72	4	9	13
10	RUO	MATEJ	Male	########	ruo_mate	47fe5b1cf	91-766223	46	2	8	10	6	8	14
11	Sankari	ABUL	Female	########	sankari@g	7610e2234	91-997797	5	57	97	154	76	83	159
12	Jitendra	W	Male	########	jitendra_v	8f294d3da	91-917464	94	78	57	135	6	6	12
13	Trupti	V	Female	########	trupti_v_S	3c1875805	91-864257	33	8	93	101	25	3	28
14	BOSTON	B	Male	########	boston_b(	252be354	91-762393	5	86	26	112	1	6	7
15	Sacchidan	O	Male	9/5/2006	sacchidan	9dc6cffdb	91-834257	99	1	8	9	62	86	148
16	Tanaya	RYLEA	Female	1/5/2013	tanaya_ry	f341f9477	NA	0	0	0	0	0	0	0
17	SHERIFF	G	Male	########	g_157@gn	6f667c7d7	91-974878	2	8	47	55	29	32	61
18	Vasanta	DALEY	Female	2012-0-11	vasanta_d	0ce5f7cb4	NA	0	0	0	0	0	0	0
19	MARLOWE	JOHN-JOE	Male	1/7/2013	marlowe_	73d71067t	91-931812	4	4	78	82	9	78	87
20	DILLON	L	Male	########	dillon_891	3abd011e	91-761196	5	5	69	74	89	7	96
21	Sushma	Q	Female	########	sushma_q	39ddb459	91-943599	1	65	9	74	92	4	96
22	CHESNEY	M	Male	3/2/2010	m@gmail.	a5c80375a	91-952178	38	63	28	91	3	3	6
23	Chandanil	Palashkus	Female	3/3/2012	chandanik	4c91f172f	91-753396	57	1	52	53	5	14	19
24	Nandita	Abhay	Female	2012-4-31	nandita_a	b691dd9f5	91-837153	5	22	17	39	8	32	40
25	TAHMID	O	Male	########	o@yahoo.	eb80de33	91-774711	79	11	14	25	2	3	5

Figure-4. Sample data sets with high dimensions and attributes in data streams.

As shown in Figure-4, we are taken data sets from different social networks using API (Application Programming Interface) developed by java for extracting relational data with multi attributes. And also, we take data with relational attributes shown in Figure-5.

1	VE_TOTAL	PERSONS	PEDS	NO_LANES	SP_LIMIT	FATALS	DRUNK_DR
2	1	2	1	3	0	1	0
3	1	1	0	2	0	1	1
4	3	3	0	2	0	1	0
5	2	2	0	2	0	1	0
6	1	6	0	2	0	3	0
7	2	8	0	2	15	1	0
8	1	2	1	2	15	1	1
9	1	2	1	2	15	1	2
10	1	1	0	2	20	1	1
11	1	2	0	2	20	1	0
12	1	4	0	2	20	1	0
13	1	1	0	9	25	1	1
14	1	3	2	6	25	1	0
15	2	4	0	5	25	1	1
16	2	6	0	4	25	1	1
17	1	4	0	4	25	2	0
18	1	2	1	4	25	1	0
19	3	4	0	4	25	1	1
20	1	2	1	4	25	1	0
21	1	2	1	4	25	1	2
22	4	6	1	4	25	1	0
23	2	6	0	4	25	1	0
24	1	7	1	4	25	1	0
25	3	3	0	4	25	1	0

Figure-5. Accident data with different attributes.

We will take accident data from different areas in recent contribution of relational attributes as shown in Figure-5.

4.1 Implementation procedure: Implementation of MAMG with different multi attributes follows following procedure for data processing. To handle data points with clustering results in real time data instances, different level of expressive data is processed with different durable places $\left[V_{L_{C(i,\beta)}},\ U_{L_{C(i,\beta)}}\right]$ for the mean as $L_{C(i,\beta)}$ with feasible validity C as follows:

$$V_{L_{C(i,\beta)}} = L_{X(i,\beta)} - 2.34\frac{S_{X(i,\beta)}}{\sqrt{n}} \tag{1}$$

$$U_{L_{C(i,\beta)}} = L_{V(i,\beta)} + 2.34\frac{S_{V(i,\beta)}}{\sqrt{n}} \tag{2}$$

Shown in Figure-5 $S_{X(i,\beta)}$ is Standard Deviation (SD) of different index values across different group labels clustering methods I with data set β. Compare proposed approach with existing approaches in this region, proposed approach gives better performance from different class instances with clusters.

$$B_{X(i)} = \sum_{\forall\beta\in DT}\ \sum_{\forall i^*\in CN, i^*\neq i} better_C^{\beta}(i,i^*), \tag{3}$$

$$better_C^{\beta}(i,i^*) = \{\begin{matrix} 1 & ifL_{XC(i,\beta)} > U_{XC(i^*,\beta)} \\ 0 & otherwise \end{matrix}\} \tag{4}$$

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Likely, in this region no. of approaches or techniques our method i<CM is simultaneously with different competitors, WC(i) in validity presentation centre presentation C to be calculated as follows:

$$W_{C(i)} = \sum_{\forall \beta \in DT} \sum_{\forall i^* \in CM, i^* \neq i} worse_C^{\beta}(i, i^*), \tag{5}$$

$$worse_C^{\beta}(i, i^*) = \{ \begin{matrix} 1 & if U_{XC(i^*, \beta)} < L_{XC(i,\beta)} \\ 0 & otherwise \end{matrix} \} \tag{6}$$

By using these equations from 1-6, gives implementation definitions for multi attributes for data processing. The efficiency of MAMG to define multi features in real time attributes.

4.2 Experimental Results: As shown in Table-1, proposed approach gives accuracy of real world entity in different data sets like Accident, diabetes, Economy Ratings, Student marks performance in different formations. Sample results for after performing proposed approach for different data attributes shown in Figure-6.

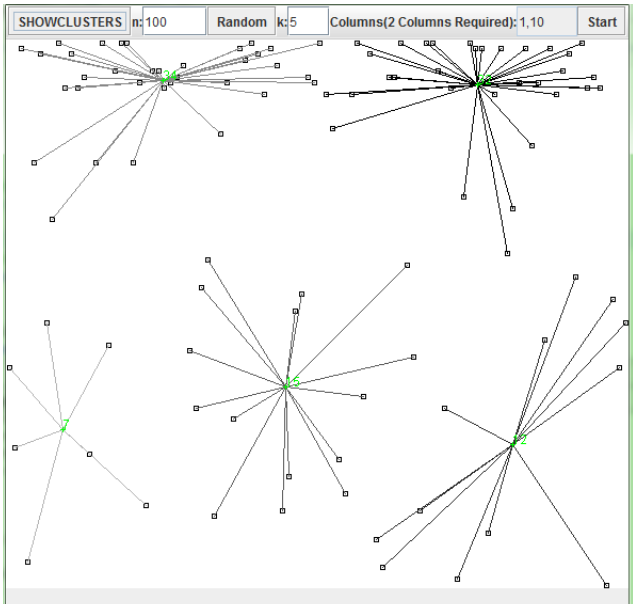

Figure-6. Multi attribute data representation with different data items.

As shown in Figure-6, first upload data sets to our proposed approach for attribute selection with different relations and then process each attribute as data point, then perform Euclidian distances between attributes for efficient data classification with multi attributes. Cluster formation based on different attributes with distance calculation shown in Figure-7.

```
--->Centroid (442,42) Aligning with (1085.0,18.0) @ (493,99)
--->Centroid (442,42) Aligning with (1086.0,7.0) @ (499,40)
--->Centroid (475,379) Aligning with (1087.0,52.0) @ (505,281)
--->Centroid (442,42) Aligning with (1088.0,1.0) @ (510,8)
--->Centroid (442,42) Aligning with (1089.0,9.0) @ (516,51)
--->Centroid (442,42) Aligning with (1090.0,5.0) @ (522,29)
--->Centroid (442,42) Aligning with (1091.0,2.0) @ (528,13)
--->Centroid (475,379) Aligning with (1092.0,41.0) @ (534,222)
--->Centroid (442,42) Aligning with (1093.0,0.0) @ (540,3)
--->Centroid (442,42) Aligning with (1094.0,8.0) @ (545,45)
--->Centroid (442,42) Aligning with (1095.0,6.0) @ (551,35)
--->Centroid (442,42) Aligning with (1096.0,8.0) @ (557,45)
--->Centroid (475,379) Aligning with (1097.0,95.0) @ (563,511)
--->Centroid (475,379) Aligning with (1098.0,45.0) @ (569,243)
--->Centroid (475,379) Aligning with (1099.0,57.0) @ (575,307)
--->Centroid (475,379) Aligning with (1100.0,49.0) @ (581,265)
-----------------------------------------------------------

Edge Cluster Walk

[225,Cluster : 1 @ (442,42) clusters count are : 32]
                        [174,Cluster : 4 @ (147,38) clusters count are : 34]
                [123,Cluster : 3 @ (262,325) clusters count are : 15]
        [62,Cluster : 2 @ (475,379) clusters count are : 12]
                [19,Cluster : 5 @ (53,365) clusters count are : 7]
Cluster Ensemble runs in 1.039 secs
```

Figure-7. Cluster formation for different edges for grouping multi attributes.

Time efficiency for our proposed approach shown in Figure-8, different data sets like accident, diabetes, with multi attributes in recent feature selection with randomly progress real time data streams.

VOL. 13, NO. 9, MAY2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

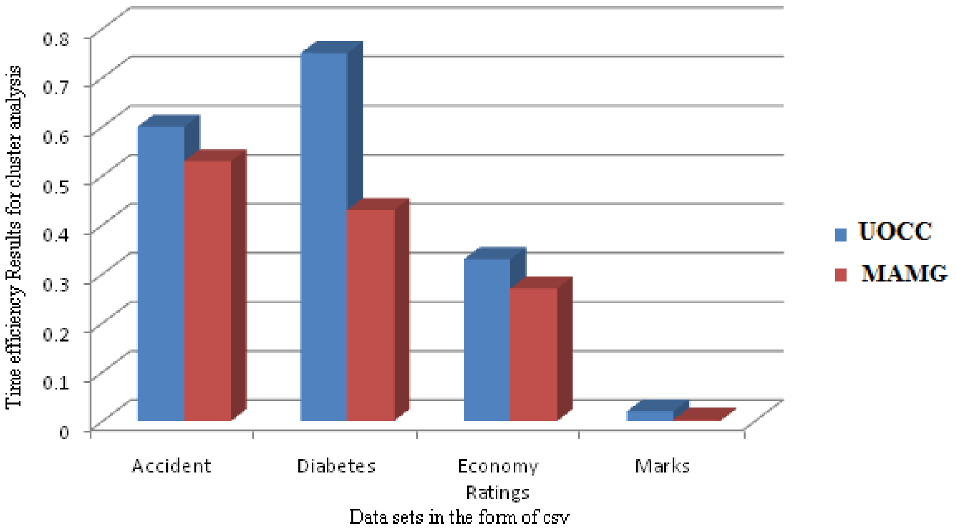

Figure-8. Results comparison with different labels in terms of time efficiency in different data sets.

Furthermore, MAMG works constantly better than its competitors with all different selection measurements, while UOCC gives least performance on class instances. Realize that a bigger selection outcomes in an enhanced perfection with better time efficiency results representation.

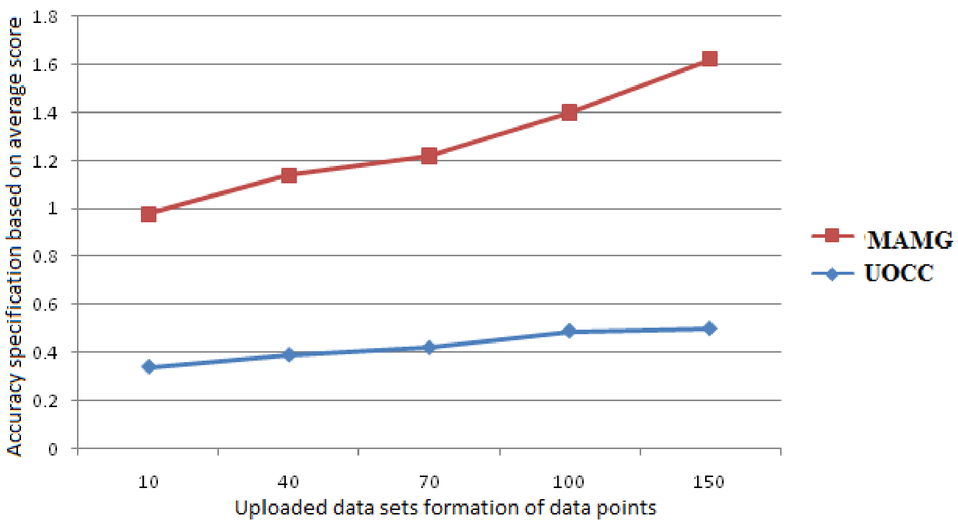

Figure-9. Accuracy evaluation analysis of proposed approach and existing approach for cluster formation.

These results are appear with MAGM with different multi objects. To start with, we make the limited score to capture the area uncertainty in light of every illustration's close by information perform, and after that produce a doubtful UOCC classifier by combining unrelated or uncertain data into a MAGM with Support Vector Machine (SVM)-based studying collaborative representation. We extend SVM technique to summarize the understanding of the consumer with replicable class instances. Wide assessments have revealed that our unverifiable one category studying can get an excellent experiment evaluation and is less sensitive to fuss in connection with the standard one-class SVM. The assessments additionally illustrate that the support vectors-based collection technique can well reduce the understanding of the consumer in connection with emphasize centered collection way of concept summary learning.

5. SUMMARY

We propose and develop Multi Attribute Grouping Method for data exploration in different data sources with multi object orientation in cluster relational data bases. This paper presents novel MAGM to categorize data based on different attributes from multi-dimensional data sources. It constructs and transforms matrix formation into attribute partition based on graph procedure. Our experimental results give efficient and

ARPN Journal of Engineering and Applied Sciences

effective approaches to configure data sets to measure attributes and combine those attributes using link based methodology It gives effective results in multi attribute combination from cluster relational data source with semantic data structure with similarity measures with feature partition. Our future work relates and extends to detect data redundancy in categorical data based on multiple attributes.

REFERENCES

[1] Bo Liu, Yanshan Xiao, Philip S. Yu. 2014. Uncertain One-Class Learning and Concept Summarization Learning on Uncertain Data Streams. IEEE Transactions on Knowledge and Data Engineering. 26(2).

[2] C.C. Aggarwal, Y. Xie and P.S. Yu. 2011. On Dynamic Data-Driven Selection of Sensor Streams. Proc. 17th ACM SIGKDD Int'l Conf.Knowledge Discovery and Data Mining (KDD). pp. 1226-1234.

[3] C.C. Aggarwal and P.S. Yu. 2009. A Survey of Uncertain Data Algorithms and Applications. IEEE Trans. Knowledge and Data Eng. 21(5): 609-623.

[4] F. Bonchi, M.V. Leeuwen and A. Ukkonen. 2011. Characterizing Uncertain Data Using Compression. Proc. SIAM Conf. DataMining. pp. 534-545.

[5] F. Bovoloa, G. Camps-Vallsb and L. Bruzzonea. 2010. A Support Vector Domain Method for Change Detection in Multitemporal Images. Pattern Recognition Letters. 31(10): 1148-1154.

[6] L. Chen and C. Wang. 2010. Continuous Subgraph Pattern Search over Certain and Uncertain Graph Streams. IEEE Trans. Knowledge and Data Eng. 22(8): 1093-1109.

[7] B. Liu, Y. Xiao, L. Cao and P.S. Yu. 2010. Vote-Based LELC for Positive and Unlabeled Textual Data Streams. Proc. IEEE Int'l Conf. Data Mining Workshops (ICDM), pp. 951-958, 2010.

[8] R. Murthy, R. Ikeda and J. Widom. 2011. Making Aggregation Work in Uncertain and Probabilistic Databases. IEEE Trans. Knowledge and Data Eng. 22(8): 1261-1273.

[9] L. Sun, R. Cheng, D.W. Cheung and J. Cheng. 2010. Mining Uncertain Data with Probabilistic Guarantees. Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining. pp. 273-282.

[10] M. Takruri, S. Rajasegarar, S. Challa, C. Leckie and M. Palaniswami. 2011. Spatio-Temporal Modelling-Based Drift-Aware Wireless Sensor Networks. Wireless Sensor Systems. 1(2): 110-122.

[11] S. Tsang, B. Kao, K.Y. Yip, W.S. Ho and S.D. Lee. 2011. Decision Trees for Uncertain Data. IEEE Trans. Knowledge and Data Eng. 23(1): 64-78.

[12] L. Trung, T. Dat, N. Phuoc, M. Wanli and D. Sharma. 2011. Multiple Distribution Data Description Learning Method for Novelty Detection. Proc. Int'l Joint Conf. Neural Networks (IJCNN). pp. 2321-2326.

[13] S.M. Yuen, Y. Tao, X. Xiao, J. Pei and D. Zhang. 2010. Superseding Nearest Neighbor Search on Uncertain Spatial Databases. IEEE Trans. Knowledge and Data Eng. 22(7): 1041-1055.

[14] X. Zhu, W. Ding, P.S. Yu and C. Zhang. 2011. One-Class Learning and Concept Summarization for Data Streams. Knowledge and Information Systems. 28(3): 523-553.

[15] Z. Zou, H. Gao and J. Li. 2010. Discovering Frequent Subgraphs over Uncertain Graph Databases under Probabilistic Semantics. Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining. pp. 633-642.

[16] Natthakan Iam-On, Tossapon Boongoen, Simon Garrett and Chris Price. 2012. Link-Based Cluster Ensemble Approach for Categorical Data Clustering. IEEE Transactions On Knowledge and Data Engineering. 24(3).

[17] T. Boongoen, Q. Shen, and C. Price. 2010. Disclosing False Identity through Hybrid Link Analysis. Artificial Intelligence and Law. 18(1): 77-102.