IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

A Survey on Incremental Attribute Reduction Method for Dynamic Data mining Decision Systems

D. Ramana Kumar[1*], S. Krishna Mohan Rao[2], K. Rajeshwar Rao[3]

[1]Department of Computer Science and Engineering, JNTUH, Hyderabad, India
[2]GIET, Bhubaneswar, Odisha, India
[3]Department of Computer Science and Engineering, SIET, Hyderabad, India

Available online at: www.ijcseonline.org

Accepted: 13/Dec/2018, Published: 31/Dec/2018

Abstract- In dynamic data mining situations, the attribute decrease issue has three issues: variety of protest sets, variety of trait sets and variety of property estimations. For the initial two issues, a couple of accomplishments have been made. For variety of the property estimations, current characteristic decrease approaches are not productive, in light of the fact that the strategy turns into a non-incremental or wasteful one sometimes. With the end goal to address this, we initially present the idea of an irregularity degree in a deficient choice framework and demonstrate that the property decrease dependent on the irregularity degree is proportional to that dependent on the positive area. At that point, three refresh procedures of irregularity degree for dynamic fragmented choice frameworks are given. At long last, the system of the incremental attribute decrease calculation is proposed.

Keywords- DIDS, mechanism in DIDS

I. INTRODUCTION

As of late, numerous researchers have gotten significant accomplishments in the fields of machine learning, man-made brainpower, design acknowledgment, information mining [4]. Characteristic decrease is one of the examination problem areas in harsh set hypothesis [8]. The primary thought is to choose a few qualities that hold the characterization capacity. Be that as it may, with the fast improvement of data innovation, the world has just entered a rapid data. As the information from each field changes after some time, the measure of information that should be recorded and handled is expanding. At long last, this information will turn into a colossal measure of dynamic complex information. Step by step instructions to store, utilize and uncover the various entangled information is one of the vital issues that should be illuminated with PC innovation and data handling capacity. In the interim, these information could be inadequate because of mistakes caused by oversights or insufficiency of data, credited to absence of gathering or other human components. These dynamic and deficient information are regularly experienced, all things considered. For instance, when a robot is on a mission, it can see a great deal of highlight data in nature. The data is fragmented and changes with time, we have to get the valuable component data from this dynamic complex data to enable a robot to perform errands consummately. For another precedent, in stock investigation, information changes quickly. With the end goal to get okay basic leadership, we should discover important information from the dynamic information. These two issues are the quality decrease for dynamic and fragmented information. Therefore, for both fragmented and dynamic information, the quality decrease issue has certain hugeness, which gives an unpleasant set strategy to secure the learning from complex information. At present, inquire about on quality decrease of dynamic information for the most part centers around three perspectives: variety of protest sets [12], variety of property sets and variety of characteristic qualities. For research on the variety of protest sets, Chen etal. proposed incremental strategies for refreshing approximations under a variable accuracy unpleasant set model where the data framework is refreshed by embeddings or erasing a protest. Fan etal. [15]proposed a dynamic trait decrease calculation, where another protest is added to the question set. Liang etal. [16]proposed a gathering incremental trait decrease calculation, where a gathering of items are included. Shu etal. [17]presented a positive locale refresh component as for the including and erasing of protest sets in fragmented choice frameworks and a dynamic trait decrease calculation is proposed. Jing etal. [18]proposed an incremental property decrease approach dependent on learning granularity where a few items shift progressively. For research on the variety of quality sets, Wang etal. built up a measurement incremental calculation for dynamic informational collections dependent on data entropy and the key advance of the improvement is the refresh component of data entropy. Shu etal. displayed an incremental trait decrease calculation dependent on inadequate choice framework, or, in other words the investigation of the refresh component of the positive locale on account of the difference in the quality set. Jing etal. utilized the perceptibility network technique to refresh the properties of learning granularity, at that point proposed an incremental quality decrease dependent on the ascertained information granularity when numerous ascribes are added to the choice framework. Considering items and characteristics fluctuate all the while, Chen etal. proposed a calculation for refreshing approximations in choice theoretic unpleasant sets for dynamic information mining by utilizing equality highlight frameworks. Wang etal. proposed a novel incremental disentangled calculation that can productively refresh approximations of a predominance based harsh set methodology when protests and traits increment all the while. For the exploration on variety of characteristic qualities, Wang etal. displayed the refresh system of data entropy on account of progressively differing verifiable information, and proposed an incremental characteristic decrease. Shu etal. displayed the refresh instrument of the positive district where the trait estimations of different articles fluctuate at the same time, and proposed an incremental property decrease dependent on the positive locale. Li etal. proposed an incremental refresh of approximations in a predominance based harsh set methodology under the variety of characteristic qualities.

II. DYNAMIC INCOMPLETE DECISION SYSTEM (DIDS) WITH VARIATION OF ATTRIBUTE VALUES

In the real world, data usually changes over time, these changes include increases, decreases, and changes of the original data, i.e. some object sets are increased or decreased, some attribute sets are increased or decreased, and some $\exists x \in U$, $\exists a \in C \cup D$and the values of f(x,a)are changed. This paper centers around the last case and talks about the quality decrease issue in powerful fragmented choice frameworks with variety of characteristic qualities. At first, we present the dynamic deficient choice framework with variety of quality qualities.

III. UPDATE MECHANISM IN DIDS

In a heuristic property decrease calculation dependent on the irregularity degree is given. The irregularity degree can be acquired by figuring the resilience class. In this manner, the key issue that ought to be tackled is: how to refresh the resilience class? In the event that we just erase unique information or include new information, the refresh component of the resilience class contains four sorts of conditions: erasing the first protest set, including the new question set, erasing the first characteristic and including the new property set. The refresh recipe of the resistance class just erases or includes information in the first choice framework. In the event that a few estimations of f(x,a)are changed, the resilience class must be changed. As we probably am aware, most ebb and flow inquire about considers these progressions as a few items' trait esteems changed. To handle this issue, we can erase the progressions identified with the first protest set and afterward include the new question set. In view of this, the protest related refresh technique for resilience class can be effortlessly introduced. The refresh equation of the resistance class just erases or includes information in the first choice framework. In the event that a few estimations of f(x,a)are changed, the resilience class must be changed. As we probably am aware, most momentum inquire about considers these progressions as a few articles' characteristic qualities changed. To handle this issue, we can erase the progressions identified with the first question set and afterward include the new protest set. In view of this, the question related refresh system for resistance class can be effortlessly displayed.

initialize $T_B^{S_U^{N+1}}(x) \leftarrow T_B^{S_U^N}(x)$;
for *each* $a_j \in U - U_{ALT}$ **do**
  $T_B^{S_U^N}(x_j) \leftarrow T_B^{S_U^N}(x_j) - U_{ALT}$
end
for *each* $x_i \in U_{ALT}$ **do**
  $T_B^{S_U^N}(x_i) \leftarrow x_i$
end
for *each* $x_k \in U - U_{ALT}$ *and* $x_i^* \in U_{ALT}$ **do**
  if $(x_k, x_i^*) \in T(B)$ **then**
    $T_B^{S_U^{N+1}}(x_k) \leftarrow T_B^{S_U^{N+1}}(x_k) \cup \{x_i^*\}$;
    $T_B^{S_U^{N+1}}(x_i^*) \leftarrow T_B^{S_U^{N+1}}(x_i^*) \cup \{x_k\}$
    unrelated object or related
  end
end
return $T_B^{S_U^{N+1}}(x)$

Fig.1. Object-related update strategy.

The protest related refresh technique is a proficiency strategy. On the off chance that a property ahas a record mistake, and all items are included. Under this condition, the question related refresh methodology must erase every single unique protest and include every single new question. This procedure is exceptionally unpredictable, another refresh technique is required for this circumstance. Like the question related procedure, we can see the progressions as a few characteristics' protest esteems changed, at that point erase the related unique quality set and include another property set.

IV. THE SYSTEM OF THE INCREMENTAL CHARACTERISTIC DECREASE ALGORITHM

The refresh instrument of DIDS is proposed. In light of this, with the end goal to quicken the characteristic decrease process, we will propose a structure of the incremental quality decrease calculation for inadequate choice framework with variety of trait esteems in the accompanying. In this structure, the first characteristic decrease result is one of sources of info, or, in other words locate the new outcome after the first information changed. As indicated by the correlation of related properties' irregularity degree when transforming, we introduce the new characteristic decrease results from the first property decrease result. Thusly, we skirt the way toward registering center characteristics.

```
for each a_ALT ∈ ΔC do
    compute ∂^{S^{N+1}}(a_ALT);
    if a_ALT ∈ RED_N then
        if ∂^{S^{N+1}}(a_ALT) ≤ ∂^{S^N}(a_ALT) then
            initialize RED_{N+1} ← RED_N;
        else
            initialize RED_{N+1} ← RED_N − {a_ALT};
        end
    else
        if ∂^{S^{N+1}}(a_ALT) < ∂^{S^N}(a_ALT) then
            initialize RED_{N+1} ← RED_N ∪ {a_ALT};
        else
            initialize RED_{N+1} ← RED_N;
        end
    end
end
select the update strategy to update ∂^{S^{N+1}}(RED_{N+1}) and ∂^{S^{N+1}}(C); /* inconsistency degree */
while ∂^{S^{N+1}}(RED_{N+1}) ≠ ∂^{S^{N+1}}(C) do
    for each a ∈ C − RED_N do
        compute SGF^{outer}(a, RED_{N+1}, D), set SGF^{outer}(a', RED_{N+1}, D)
        RED_{N+1} = RED_{N+1} ∪ {a'} and update ∂^{S^{N+1}}(RED_{N+1});
    end
end
for each a ∈ RED_{N+1} do
    compute SGF^{inner}(a, RED_{N+1}, D);
    if SGF^{inner}(a, RED_{N+1}, D) = 0 then
        RED_{N+1} = RED_{N+1} − {a} and update ∂^{S^{N+1}}(RED_{N+1});
    end
end
return RED_{N+1}
```

Fig.2. The framework of the incremental attribute reduction algorithm

V. CONCLUSION

For dynamic inadequate choice frameworks with variety of characteristic qualities, the current incremental property reduction calculations are not sufficiently viable under a few conditions. To enhance it, in this paper, we gave a quality decrease calculation dependent on the irregularity degree in fragmented choice frameworks. As per this trait reduction calculation, three methodologies for refreshing resilience classes are proposed and the refresh component of the irregularity degree is displayed. On the establishment of the refresh component, the structure of the incremental property decrease calculation in powerful inadequate choice frameworks with variety of quality qualities is proposed. A progression of investigations are led to assess the execution of the property decrease calculation dependent on the irregularity degree and the proposed incremental trait decrease calculations by utilizing diverse refresh methodologies. The outcomes showed that the characteristic decrease calculation dependent on the irregularity degree is effective and doable. Contrasted and non-incremental calculations and the cutting edge include choice calculations, the proposed incremental calculations have diverse efficiency in various circumstances. The current incremental calculations in unique fragmented choice frameworks with variety of trait esteems just contain question related system, however our methodology contains three procedures for various circumstances. Future work is to stretch out this strategy to other unpleasant set models and enhance our way to deal with fit blended complex information or enormous information.

REFERENCES

[1]R.W. Swiniarski, A. Skowron, Rough set methods in feature selection and recognition, Pattern Recognit. Lett. 24 (2003) 833–849.

[2]M. Moshkov, B. Zielosko, Combinatorial Machine Learning: A Rough Set Approach, Studies in Computational Intelligence, vol.360, Springer, Berlin, 2011.

[3]Z. Pawlak, A. Skowron, Rudiments of rough sets, Inf. Sci. 177 (2007) 3–27

[4]Z. Pawlak, Rough set, Int. J. Comput. Inf. Sci. 11 (1982) 341–356.

[5]W. Ziarko, Variable precision rough set model, J. Comput. Syst. Sci. 46 (1993) 39–59.

[6]Z. Pawlak, Rough set theory and its application to data analysis, Cybern. Syst. 29 (1998) 661–668.

[7]A.K. Jain, R.P.W. Duin, J. Mao, Statistical pattern recognition: a review, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 4–37.

[8]D. Chen, C. Wang, Q. Hu, A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets, Inf. Sci. 177 (2007) 3500–3518.

[9]Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set models, Inf. Sci. 178 (2008) 3356–3373.

[10]Y. Qian, J. Liang, W. Pedrycz, Positive approximation: an accelerator for attribute reduction in rough set theory, Artif. Intell. 174 (2010) 597–618

[11]J. Stefanowski, A. Tsoukias, Incomplete information tables and rough classification, Comput. Intell. 17 (2001) 545–566

[12]F. Hu, G. Wang, H. Huang, Incremental attribute reduction based on elementary sets, in: International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Springer-Verlag, 2005, pp.185–193.

[13]J.Y. Liang, W. Wei, Y.H. Qian, An incremental approach to computation of a core based on conditional entropy, Chin. J. Syst. Eng. Theory Pract. 4 (2008) 81–89.

[14]D. Liu, T.R. Li, D. Ruan, W.L. Zou, Anincremental approach for inducing knowledge from dynamic information systems, Fundam. Inform. 94 (2009) 245–260.

[15]Y.N. Fan, C.C. Huang, C.C. Chern, Rule induction based on an incremental rough set, Expert Syst. Appl. 36 (2009) 11439–11450.

[16]J. Liang, F. Wang, C. Dang, A group incremental approach to feature selection applying rough set technique, IEEE Trans. Knowl. Data Eng. 26 (2014) 294–308.

[17]W. Shu, W. Qian, An incremental approach to attribute reduction from dynamic incomplete decision systems in rough set theory, Data Knowl. Eng. 100 (2015) 116–132.

[18]Y. Jing, T. Li, F. Fujita, et al., An incremental attribute reduction approach based on knowledge granularity with a multi-granulation view, Inf. Sci. 411 (2017) 23–38.

[19]A. Zeng, T. Li, D. Liu, A fuzzy rough set approach for incremental feature selection on hybrid information systems, Fuzzy Sets Syst. 258 (2015) 39–60.

[20]C.C. Chan, A rough set approach to attribute generalization in data mining, Inf. Sci. 107 (1998) 169–176.

[21]T.R. Li, D. Ruan, W. Geert, J. Song, Y. Xu, A rough sets based characteristic relation approach for dynamic attribute generalization in data mining, Knowl.-Based Syst. 20 (2007) 485–494.

