

Mining Significant Patterns in Diverse Data using Our Great Algorithm

Your Name and Anne M. Denton
Department of Computer Science and Operations Research
North Dakota State University
Fargo, ND
{your.name, anne.denton}@ndsu.edu

Abstract

This is the broad topic. This is what we do. We demonstrate the efficiency and effectiveness of our algorithm on these types of data.

Keywords:

1 Introduction

The topic is of great current relevance because ... it is related to ... market basket research [1]

Fig. 1 illustrates the problem of interest. ...

2 Related Work

3 Our Great Algorithm

3.1 Relevant Data Types

We consider .. D continuous attributes, $x_i \in \mathbb{R}, 0 \leq i < D$,

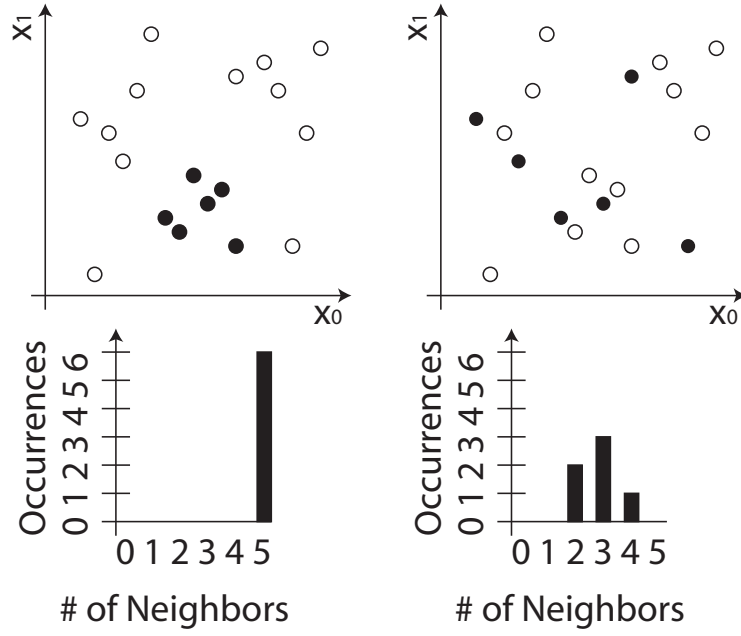


Figure 1: Schematic of This is represented as circles, that as crosses, the other thing by a dashed line. The left panel shows this and the right panel the other thing.

3.2 Outline of the Algorithm

3.3 Main Concepts

3.3.1 Central Concept

A one-line equation

$$R^{(i)} = \sigma_{B^{(i)}}(R) \quad (1)$$

an equation that extends over multiple lines

$$\begin{aligned} c_{\text{uniform}}(\mathbf{x}, \mathbf{y}) &= \begin{cases} 1 & \text{if } |\mathbf{x} - \mathbf{y}| < d/2 \\ 0 & \text{else} \end{cases} \\ &= \theta\left(\frac{d}{2} - |\mathbf{x} - \mathbf{y}|\right) \end{aligned} \quad (2)$$

Algorithm 1: Density Histogram Algorithm

```
Data: pts; /* nPts data points */
Data: items;
Result: significance; /* of each item */
1 normPts = normalize(pts);
2 hist = zeros(1,nPts); /* vector of zeros */
3 foreach it  $\in$  items do
4   itemPts = findPoints(normPts,it);
5   foreach x  $\in$  itemPts do
6     dens = NumberOfNeighbors(x);
7     hist(dens)++;
8   randHist = zeros(1, nPts);
9   for i=1:nAv do
10    randPts = randSubset(normPts,occurrences(it));
11    foreach x  $\in$  randPts do
12      dens = NumberOfNeighbors(x);
13      randHist(dens)++;
14    randHist/ = nAv;
15    significance(it) = chiSquaredGoodnessOfFit(hist, randHist);
16 return significance
```

3.3.2 Theoretical Analysis

3.4 Summary of Algorithm

Algorithm 1 describes the randomization variant of the algorithm....

4 Experimental Evaluation

The algorithm is implemented in ...

4.1 Comparison Algorithm

4.2 Evaluation on This Data

Table 1 summarizes the properties of the four data sets from separate experiments ...

Table 1: Expression Data Sets

Name	Abbr.	No. of Att.	No. of Genes
Alpha	Alp	18	6177
Cdc15	C15	24	5995
Cdc28	C28	17	6147
Elu	Elu	14	6075
—	All	73	6178

4.3 Evaluation on That Data

4.4 Evaluation on The Other Data

4.4.1 Data Sets

4.4.2 Effectiveness

4.4.3 Choice of Parameters

4.5 Performance

5 Conclusions

This is what we did and these are all our great achievements.

6 Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. IDM-0415190 and ...

References

- [1] R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In *Proc. ACM SIGMOD Int'l Conf. on Management of Data*, pages 207–216,

Washington, D.C., 26–28 1993.