

# Interestingness Measures for Fuzzy Association Rules

Attila Gyenesei and Jukka Teuhola

Turku Centre for Computer Science (TUCS)  
University of Turku, Department of Computer Science  
Lemminkäisenkatu 14, FIN-20520 Turku, Finland  
{gyenesei,teuhola}@cs.utu.fi

**Abstract.** Data mining tries to discover interesting and surprising patterns among a given data set. An important task is to develop effective measures of interestingness for evaluating and ranking the discovered patterns. A good measure should give a high rank to patterns, which have strong evidence among data, but which yet are not too obvious. Thereby the initial set of patterns can be pruned before human inspection. In this paper we study interestingness measures for generalized quantitative association rules, where the attribute domains can be fuzzy. Several interestingness measures have been developed for the discrete case, and it turns out that many of them can be generalized to fuzzy association rules, as well. More precisely, our goal is to compare the fuzzy version of confidence to some other measures, which are based on statistics and information theory. Our experiments show that although the rankings of rules are relatively similar for most of the methods, also some anomalies occur. Our suggestion is that the information-theoretic measures are a good choice when estimating the interestingness of rules, both for fuzzy and non-fuzzy domains.

## 1 Introduction

Data mining, also referred to as knowledge discovery in databases, is concerned with the nontrivial extraction of implicit, previously unknown, and potentially useful information from data [14]. One major application domain of data mining is the analysis of transactional data. The problem of mining boolean association rules over basket data was first introduced in [1], and later broadened in [2], for the case of databases consisting of categorical attributes alone.

For example, in a database maintained by a supermarket, an association rule might be of the form:

“beer and potato chips  $\rightarrow$  diapers (support: 2%, confidence: 73%)”,

which means that 2% of all database transactions contain the data items beer, potato chips and diapers, and 73% of the transactions that have the items “beer” and “potato chips” also have the item “diapers” in them. The two percentage values are referred to as *support* and *confidence*, respectively.

In practice the information in many, if not most, databases is not limited to categorical attributes (e.g. zip code, make of car), but also contains much quantitative data (e.g. age, income). The problem of mining quantitative association rules was introduced and an algorithm proposed in [17]. The algorithm involves discretizing the domains of quantitative attributes into intervals in order to reduce the domain into a categorical one. An example of such an association might be “10% of married people between 50 and 70 have at least 2 cars”.

Without a priori knowledge, however, determining the right intervals can be a tricky and difficult task due to the “catch-22” situation, as called in [17], because of the effects of *small support* and *small confidence*. Moreover, these intervals may not be concise and meaningful enough for human experts to easily obtain nontrivial knowledge from those rules discovered.

Instead of using sharp intervals, fuzzy sets were suggested in [12] to represent intervals with non-sharp boundaries. The obtained rules are called fuzzy association rules. If meaningful linguistic terms are assigned to fuzzy sets, the fuzzy association rule is more understandable. The above example could be rephrased e.g. “10% of married old people have several cars”. Algorithms for mining fuzzy association rules were proposed in ([9], [7]), but the problem is that an expert must provide the required fuzzy sets of the quantitative attributes and their corresponding membership functions. It is unrealistic to assume that experts can always provide the best fuzzy sets for fuzzy association rule mining. In [8], we tackled this problem and proposed a method to find the fuzzy sets for quantitative attributes by using clustering techniques.

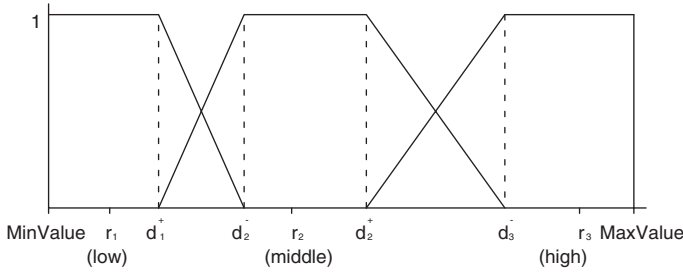
It has been recognized that a discovery system can generate a large number of patterns, most of which are of no interest to the user. To be able to prune them, researchers have defined various measures of interestingness for patterns. The most popular are confidence and support [1], others include e.g. variance and chi-square (correlation) [13], entropy gain [13], laplace [5], and intensity of implication [4]. Properties of various measures were analyzed in [3]. An extensive survey of recently proposed interestingness measures is given in [10].

The term ‘interestingness’ is often used in a subjective sense, meaning the same as ‘surprisingness’. Here we take the view that it should be also measurable in more precise terms. Although many interestingness measures have been developed for “discrete” domains, they are not directly applicable to other problem domains, such as fuzzy association rules. In this paper, we introduce generalizations of interestingness measures for fuzzy association rules, based on statistics and information theory. Especially, we present two new measures using the entropy concept. Our suggestion is that these information-theoretic measures are a good choice when estimating the interestingness of rules, both for fuzzy and non-fuzzy domains.

The rest of this paper is organized as follows. In the next section, we give a short summary of fuzzy association rules. Then we propose six measures for this fuzzy approach in Sect. 3. In Sect. 4 the experimental results are reported, comparing the proposed fuzzy interestingness measures. The paper ends with a brief conclusion in Sect. 5.

## 2 Fuzzy Association Rules

Let  $I = \{i_1, i_2, \dots, i_m\}$  be the complete set of attributes where each  $i_j$  ( $1 \leq j \leq m$ ) denotes a categorical or quantitative attribute. Note that categories are a special case of quantitative attributes, and can be handled similarly. In [8], we proposed a method to find the fuzzy sets for each quantitative attribute by using clustering techniques. We defined the goodness index  $G$  for clustering scheme evaluation, based on two criteria: compactness and separation. The clustering process determines both the number ( $c$ ) and centers ( $r_i, i = 1, \dots, c$ ) of clusters. We divide the attribute interval into  $c$  sub-intervals around the cluster centers, with a coverage of  $p$  percent between two adjacent ones, and give each subinterval a symbolic name related to its position (Fig. 1). The non-fuzzy partitioning is obtained as a special case by setting  $p$  to zero.



**Fig. 1.** Example of the proposed fuzzy partitions

For fuzzy set  $i$ ,  $d_i^+$  means the effective upper bound, and is given by:

$$d_i^+ = r_i + \frac{1}{100} \cdot \left( \frac{100 - p}{2} \right) \cdot (r_{i+1} - r_i),$$

where  $p$  is the overlap parameter in %, and  $r_i$  is the center of cluster  $i$ ,  $i = 1, 2, \dots, c - 1$ .

Similarly, for fuzzy set  $j$ ,  $d_j^-$  means the effective lower bound, given by:

$$d_j^- = r_j - \frac{1}{100} \cdot \left( \frac{100 - p}{2} \right) \cdot (r_j - r_{j-1}),$$

where  $j = 2, 3, \dots, c$ .

Then, we generate the corresponding membership function for each fuzzy set of a quantitative attribute; for formulas, see [8]. Finally, a new transformed (fuzzy) database  $D^T$  is generated from the original database by applying the discovered fuzzy sets and the membership values.

Given a database  $D^T = \{t^1, t^2, \dots, t^n\}$  with attributes  $I$  and the fuzzy sets  $F(i_j)$  associated with attributes  $i_j$  in  $I$ , we use the following form for a fuzzy association rule [12]:

**If**  $X = \{x_1, x_2, \dots, x_p\}$  **is**  $A = \{a_1, a_2, \dots, a_p\}$   
**then**  $Y = \{y_1, y_2, \dots, y_q\}$  **is**  $B = \{b_1, b_2, \dots, b_q\}$ ,

where  $a_i \in F(x_i)$ ,  $i = 1, \dots, p$ ,  $b_j \in F(y_j)$ ,  $j = 1, \dots, q$ .  $X$  and  $Y$  are ordered subsets of  $I$  and they are disjoint i.e. they share no common attributes.  $A$  and  $B$  contain the fuzzy sets associated with the corresponding attributes in  $X$  and  $Y$ . As in the binary association rule, “ $X$  is  $A$ ” is called the *antecedent* of the rule while “ $Y$  is  $B$ ” is called the *consequent* of the rule. We also denote  $Z = X \cup Y = \{z_1, \dots, z_{p+q}\}$  and  $C = A \cup B = \{c_1, \dots, c_{p+q}\}$ .

### 3 Interestingness Measures for Fuzzy Association Rules

One problem area in knowledge discovery is the development of interestingness measures for ranking the usefulness and utility of discovered patterns and rules. In this section, we first describe the fuzzy itemset measures, then we propose six other candidate measures for fuzzy association rules. Three of these six are based on traditional statistics, and their non-fuzzy counterparts have occurred many times in the literature. The three other measures, on the other hand, have their basis in the information theory, originally developed by Shannon, see e.g. [15]. Although our experiments in Sect. 4 do not show a big difference between the two categories of methods, we conjecture that the information-theoretic measures are in some cases better in capturing dependences among data. Another classification of measures would be on the basis of linear/nominal scale, but our formulations are such that this separation need not be explicit.

#### 3.1 Fuzzy Itemset Measures

Let  $D^T = \{t^1, t^2, \dots, t^n\}$  be a database, where  $n$  denotes the total number of records (‘transactions’). Let  $(Z, C)$  be an attribute-fuzzy set pair, where  $Z$  is an ordered set of attributes  $z_j$  and  $C$  is a corresponding set of fuzzy sets  $c_j$ . (From now on, we prefer to use the word “itemset” instead of “attribute-fuzzy set pair” for  $(Z, C)$  elements). If a fuzzy association rule  $(X, A) \rightarrow (Y, B)$  is interesting, it should have enough fuzzy support  $FS_{(Z, C)}$  and a high fuzzy confidence value  $FC_{((X, A), (Y, B))}$ , where  $Z = X \cup Y$ ,  $C = A \cup B$ .

The fuzzy support value is calculated by multiplying the membership grade of each  $(z_j, c_j)$ , summing them, then dividing the sum by the number of records [12]. We prefer the product operator as the fuzzy AND, instead of the normal minimum, because it better distinguishes high- and low-support transactions.

$$FS_{(Z, C)} = \frac{\sum_{i=1}^n \prod_{j=1}^m (t^i [(z_j, c_j)])}{n},$$

where  $m$  is the number of items in itemset  $(Z, C)$ .

The fuzzy confidence value is calculated as follows:

$$FC_{((X, A), (Y, B))} = \frac{FS_{(Z, C)}}{FS_{(X, A)}}.$$

**Table 1.** Part of database containing fuzzy membership

...	$(Balance, low)$	...	$(Income, low)$	...	$(Credit, low)$	$(Credit, high)$	...
	0.2		0.2		0.1	0.9	
	0.4		0.2		0.1	0.9	
...	0.9	...	0.8	...	0.7	0.3	...
	0.9		0.8		0.9	0.1	
	0.6		0.4		0.9	0.1	

Both of the above formulas are direct generalizations of the corresponding formulas for the non-fuzzy case [1].

We shall use the following two rules to demonstrate the calculation of interestingness measures. The data behind the rules are presented in Table 1.

Let  $X = \{Balance, Income\}$ ,  $A = \{low, low\}$ ,  $Y = \{Credit\}$ ,  $B = \{low\}$ , and  $\bar{B} = \{high\}$ . Rule<sub>1</sub>  $(X, A) \rightarrow (Y, B)$  is given by:

“If *Balance* is *low* and *Income* is *low* then *Credit* is *low*”,

and Rule<sub>2</sub>  $(X, A) \rightarrow (Y, \bar{B})$  is phrased as:

“If *Balance* is *low* and *Income* is *low* then *Credit* is *high*”.

The consequents are thus complements of each other. From the table data, we can easily see that Rule<sub>1</sub> should be classified as ‘valid’, whereas Rule<sub>2</sub> should not. When introducing the different measures of interestingness, we will check, how they are able to confirm this observation.

*Example 1.* Fuzzy confidence does a good job in assessing the rules:

$$FC_{((X,A),(Y,B))} = \frac{0.004 + 0.008 + 0.504 + 0.648 + 0.216}{0.04 + 0.08 + 0.72 + 0.72 + 0.24} = 0.766$$

$$FC_{((X,A),(Y,\bar{B}))} = \frac{0.036 + 0.072 + 0.216 + 0.072 + 0.024}{0.04 + 0.08 + 0.72 + 0.72 + 0.24} = 0.233$$

For Rule<sub>1</sub>,  $FS_{(Y,B)} = 0.54$  and  $FS_{(Z,C)} = 0.276$ . Similarly,  $FS_{(Y,\bar{B})} = 0.46$ , and  $FS_{(Z,C)} = 0.084$  for Rule<sub>2</sub>. In both cases,  $FS_{(X,A)} = 0.36$ .

### 3.2 Fuzzy Covariance Measure

Covariance is one of the simplest measures of dependence, based on the co-occurrence of the antecedent  $(X, A)$  and consequent  $(Y, B)$ . If they co-occur clearly more often than what can be expected in an independent case, then the rule  $(X, A) \rightarrow (Y, B)$  is potentially interesting. Piatetsky-Shapiro called this measure a *rule-interest* function [14]. We extend it to the fuzzy case, and define the covariance measure as:

$$Cov_{((X,A),(Y,B))} = FS_{(Z,C)} - FS_{(X,A)} \cdot FS_{(Y,B)}.$$

*Example 2.* The covariance measures for our sample rules are:

$$Cov_{((X,A),(Y,B))} = 0.0816, \text{ and } Cov_{((X,A),(Y,\bar{B}))} = -0.0816.$$

### 3.3 Fuzzy Correlation Measure

Covariance has generally the drawback that it does not take distributions into consideration. Therefore, in statistics, it is more common to use so called correlation measure, where this drawback has been eliminated. Again, we have to generalize the non-fuzzy formula to the fuzzy case, and obtain:

$$Corr_{((X,A),(Y,B))} = \frac{Cov_{((X,A),(Y,B))}}{\sqrt{Var_{(X,A)} \cdot Var_{(Y,B)}}},$$

where

$$Var_{(X,A)} = FS_{(X,A)}^2 - (FS_{(X,A)})^2,$$

$$FS_{(X,A)}^2 = \frac{\sum_{i=1}^n (II_{j=1}^m t^i [(x_j, a_j)])^2}{n},$$

similarly for  $(Y, B)$ .

These definitions are extensions of the basic formulas of variance and covariance. The value of the fuzzy correlation ranges from -1 to 1. Only a positive value tells that the antecedent and consequent are related. The higher the value is, the more related they are.

*Example 3.* Again, applying the formula to our two sample rules, we obtain:

$$Corr_{((X,A),(Y,B))} = 0.738, \text{ and } Corr_{((X,A),(Y,\bar{B}))} = -0.738.$$

### 3.4 Fuzzy I-Measure

As an example of a more ‘exotic’ probability-based interestingness measure, we give the fuzzy version of a so-called *I-measure* suggested by Gray and Orłowska [6]. Though it has some structural similarity with correlation, we regard it rather as a heuristic measure. The fuzzy *I-measure* is defined as:

$$I_{((X,A),(Y,B))} = \left[ \left( \frac{FS_{(Z,C)}}{FS_{(X,A)} \cdot FS_{(Y,B)}} \right)^k - 1 \right] \cdot (FS_{(X,A)} \cdot FS_{(Y,B)})^m,$$

where  $k$  and  $m$  are weight parameters of the two terms. A practical problem in applying this measure is the selection of these parameters.

*Example 4.* The *I-measure* values for our two sample rules (for  $k = m = 2$ ) are:

$$I_{((X,A),(Y,B))} = 0.038, \text{ and } I_{((X,A),(Y,\bar{B}))} = -0.02.$$

In data mining, an association rule  $X \rightarrow Y$  usually means that  $X$  implies  $Y$  and we cannot assume  $Y$  also implies  $X$ . Covariance, Correlation and *I-measure* are *symmetric* with respect to  $(X, A)$  and  $(Y, B)$ . Thus, we can use them only as *non-directed* measures of dependence.

### 3.5 Fuzzy Unconditional Entropy (UE) Measure

Assume that we want to evaluate rule  $(X, A) \rightarrow (Y, B)$ , and denote  $Z = X \cup Y$  and  $C = A \cup B$ . For  $(X, A)$ ,  $(Y, B)$ , and  $(Z, C)$  we can calculate the probability of occurrence, based on the transactions. These are in fact the same as the (fuzzy) support  $FS_{(X,A)}$ ,  $FS_{(Y,B)}$ , and  $FS_{(Z,C)}$ . If  $(X, A)$  and  $(Y, B)$  are independent, then it holds that

$$FS_{(Z,C)} = FS_{(X,A)} \cdot FS_{(Y,B)}.$$

However, if there is a dependence, then the equality does not hold. The degree of correlation can be measured as follows. We determine the information amount needed by assuming independence; we call this *independence entropy*, denoted  $H_{((X,A),(Y,B))}$  and computed as follows:

$$\begin{aligned} H_{((X,A),(Y,B))} = & -FS_{(Z,C)} \cdot \log_2 (FS_{(X,A)} \cdot FS_{(Y,B)}) - \\ & - (1 - FS_{(Z,C)}) \cdot \log_2 (1 - FS_{(X,A)} \cdot FS_{(Y,B)}). \end{aligned}$$

This represents the amount of information needed per transaction, when using a (false) assumption of independence, applied when true probability is  $FS_{(Z,C)}$ . The true entropy of  $(Z, C)$  is computed as follows:

$$H_{(Z,C)} = -FS_{(Z,C)} \cdot \log_2 (FS_{(Z,C)}) - (1 - FS_{(Z,C)}) \cdot \log_2 (1 - FS_{(Z,C)}).$$

Since this formula uses precise probabilities, its value is always smaller than or equal to the independence entropy. Moreover, their difference is larger when the dependence is higher. Therefore, we get a good measure of correlation as the difference, which we call *unconditional entropy (UE)*:

$$UE_{((X,A),(Y,B))} = H_{((X,A),(Y,B))} - H_{(Z,C)}.$$

Notice that although the measure is always non-negative, the related correlation can be either positive or negative, so that  $(X, A)$  and  $(Y, B)$  occur together either more or less frequently than in an independent case. Therefore, the true consequent of the (interesting) rule can be either  $(Y, B)$  or its complement. The latter holds if  $FS_{(Z,C)} < FS_{(X,A)} \cdot FS_{(Y,B)}$ , i.e. covariance is  $< 0$ .

*Example 5.* Let us compute the *UE*-values for our two sample rules:

$$UE_{((X,A),(Y,B))} = 0.878 - 0.850 = 0.028, \text{ and } UE_{((X,A),(Y,\bar{B}))} = 0.041.$$

Although the latter value is higher than the former, the condition  $FS_{(Z,C)} = 0.084 < FS_{(X,A)} \cdot FS_{(Y,\bar{B})} = 0.36 \cdot 0.46 = 0.1656$  holds, and we conclude that Rule<sub>2</sub> is not valid. Instead, for Rule<sub>1</sub>  $FS_{(Z,C)} = 0.276 > FS_{(X,A)} \cdot FS_{(Y,B)} = 0.1944$ , so Rule<sub>1</sub> is a ‘good’ one.

### 3.6 Fuzzy Conditional Entropy (CE) Measure

*UE*-measure is analogous to the correlation of a rule in the sense that both formulas are symmetric with respect to  $(X, A)$  and  $(Y, B)$ . We now develop another measure, which makes a distinction between the antecedent and consequent. Hereby we obtain an information-theoretic counterpart of confidence. The reasoning resembles the derivation of *UE*, but from the consequent's point of view. The unconditional entropy of  $(Y, B)$  is computed as follows:

$$H_{(Y,B)} = -FS_{(Y,B)} \cdot \log_2 (FS_{(Y,B)}) - (1 - FS_{(Y,B)}) \cdot \log_2 (1 - FS_{(Y,B)}).$$

If  $(X, A)$  affects in some way on  $(Y, B)$ , then the conditional probability  $P_{((Y,B)|(X,A))}$  is different from  $P_{(Y,B)}$ . Notice that the conditional probability is the same as (fuzzy) confidence  $FC_{((X,A),(Y,B))}$ , defined earlier. The conditional entropy is computed as

$$H_{((Y,B)|(X,A))} = -FC_{((X,A),(Y,B))} \cdot \log_2 (FC_{((X,A),(Y,B))}) - (1 - FC_{((X,A),(Y,B))}) \cdot \log_2 (1 - FC_{((X,A),(Y,B))}).$$

Since the conditional entropy uses a more precise value for the probability of  $(Y, B)$  among the subset studied (transactions satisfying  $(X, A)$ ), for 'true' rules,  $H_{(Y,B)}$  should be larger than  $H_{((Y,B)|(X,A))}$ . Their difference represents the deviation from the independent case, and measures the dependence of  $(Y, B)$  on  $(X, A)$ . The interestingness measure is thus defined as

$$CE_{((X,A),(Y,B))} = H_{(Y,B)} - H_{((Y,B)|(X,A))}.$$

The larger the value, the higher the dependence. As for *UE*, also here the actual consequent of a rule classified as *interesting* can be either  $(Y, B)$  or its complement. The latter holds if  $FC_{((X,A),(Y,B))} < FS_{(Y,B)}$ .

It should be noted that *CE* is similar to the *Theil* index [18], in the sense that both measure deviation from the expected entropy.

*Example 6.* The *CE-measure* gives the same value for both of our sample rules, because the consequents are complements to each other:

$$CE_{((X,A),(Y,B))} = CE_{((X,A),(Y,\overline{B}))} = 0.995 - 0.784 = 0.211.$$

This is just what should happen in information-theoretic sense. Our additional condition determines that Rule<sub>1</sub> is positive and Rule<sub>2</sub> is negative.

### 3.7 Fuzzy J-measure

Information theory has naturally been applied to measuring interestingness of rules before. One such measure, so called *J-measure*, was suggested by Smyth and Goodman [16]. It can be generalized to the fuzzy rules as follows:

$$J_{((X,A),(Y,B))} = FS_{(Y,B)} \cdot \left[ \frac{FS_{(Z,C)}}{FS_{(Y,B)}} \cdot \log_2 \left( \frac{FS_{(Z,C)}}{FS_{(Y,B)} \cdot FS_{(X,A)}} \right) + \right.$$



$$+ \left( 1 - \frac{FS_{(Z,C)}}{FS_{(Y,B)}} \right) \cdot \log_2 \left( \frac{1 - \frac{FS_{(Z,C)}}{FS_{(Y,B)}}}{1 - FS_{(X,A)}} \right) \Bigg]$$

The first term  $FS_{(Y,B)}$  measures the generality of the rule. The term inside the square brackets measures the relative entropy of  $(Y, B)$  on  $(X, A)$ , the similarity of two probability (support) distributions. Though the  $J$ -measure is different from our  $UE$ -measure, in experiments it gave rather similar rankings to rules (see next section).

*Example 7.* The  $J$ -measures for our two test rules are:

$$J_{((X,A),(Y,B))} = 0.037, \text{ and } J_{((X,A),(Y,\bar{B}))} = 0.050.$$

## 4 Experimental Results

We assessed the effectiveness of our interestingness measures by experimenting with a real-life dataset, which comes from a research by the U.S. Census Bureau. The data had 6 quantitative attributes. This database has been used in previous data mining research ([7], [8]) and will not be described again here.

Using support threshold = 20% and confidence threshold = 50%, we get exactly 20 rules, which are evaluated in the tests. Table 2 and Table 3 describe the calculated interestingness and the assigned ranks, respectively, as determined by the corresponding interestingness measure.

To quantify the extent of the ranking similarities between the seven measures, we computed the correlation coefficient for each pair of interestingness measures, see Table 4. The coefficients vary from a low 0.243 for the pair *Conf* and *I-measure*, to a high of 0.988 for the pair *UE*- and *J-measure*.

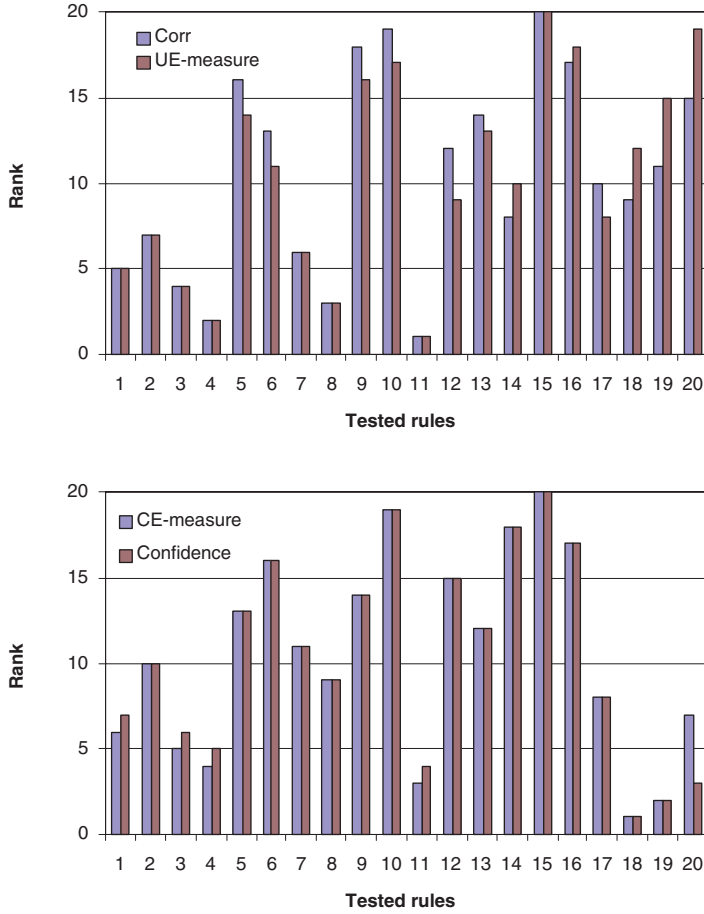
We found two distinct groups of measures, which are ranked similarly. One group consists of the non-directed measures *Cov*, *Corr*, *I-measure*, *UE-measure*, and *J-measure*. The other group consists of the directed measures *CE*, and *Conf*. However, there are no negative correlations between the two groups. Two representatives from both groups are shown in Fig. 2a and Fig. 2b.

At this point the reader may wonder, what advantage do information-theoretic measures give over statistical ones, if any. The difference comes e.g. in cases where support values are rather high. High support implies also a rather high confidence, even in a case where the antecedent and consequent are independent.

However, *CE* gives a value  $\approx 0$  in this case, pruning the rule. That will happen also with the ‘symmetric’ measures *Cov*, *Corr*, *I-measure*, and *UE-measure*, but their drawback is lack of direction. Thus, our conjecture is that *CE* is a good means of measuring the interestingness of rules.

## 5 Conclusion and Future Work

In this paper we have studied interestingness measures for generalized quantitative association rules, where the attribute domains can be fuzzy. We compared



**Fig. 2.** Rankings of interestingness measures - (a) group<sub>1</sub>, (b) group<sub>2</sub>

the fuzzy version of confidence to six other measures, three of which were statistical, and the rest three were information-theoretic, based on the entropy concept.

The experiments show that rankings of rules are relatively similar for most of the methods, but also several inversions appeared in the ranking order. Scores of interestingness measures were used to compute the correlation coefficients, revealing two categories of measures, the directed and non-directed ones. We suggest that the information-theoretic measures are a good choice when estimating the interestingness of rules, both for fuzzy and non-fuzzy domains.

Here we compared the measures of interestingness only by means of numerical rankings, obtained from experiments. In the future, we plan to compare them

**Table 2.** Interestingness measures - scores

Rule	<i>Conf</i>	<i>Cov</i>	<i>Corr</i>	<i>I-measure</i>	<i>UE-measure</i>	<i>CE-measure</i>	<i>J-measure</i>
1	0.8384	0.0874	0.4522	0.0297	0.0427	0.3619	0.0524
2	0.7841	0.0812	0.4018	0.0292	0.0347	0.2473	0.0435
3	0.8418	0.0933	0.4728	0.0335	0.0464	0.3700	0.0580
4	0.8504	0.1601	0.7232	0.0970	0.0938	0.3911	0.1558
5	0.7067	0.0442	0.2102	0.0205	0.0081	0.1066	0.0102
6	0.6416	0.0528	0.2450	0.0216	0.0128	0.0586	0.0171
7	0.7655	0.0921	0.4273	0.0395	0.0387	0.2141	0.0519
8	0.8022	0.1514	0.6704	0.0967	0.0805	0.2824	0.1425
9	0.6804	0.0398	0.1809	0.0208	0.0060	0.0756	0.0080
10	0.5768	0.0338	0.1498	0.0150	0.0049	0.0170	0.0068
11	0.8521	0.1680	0.7560	0.1064	0.1005	0.3954	0.1751
12	0.6468	0.0576	0.2635	0.0249	0.0146	0.0630	0.0199
13	0.7093	0.0475	0.2231	0.0233	0.0090	0.1100	0.0116
14	0.5918	0.0682	0.3433	0.0522	0.0144	0.0244	0.0367
15	0.5117	-0.0064	-0.0316	-0.0046	0.0001	-0.0008	0.0003
16	0.6399	0.0394	0.2035	0.0340	0.0046	0.0369	0.0097
17	0.8148	0.0569	0.2923	0.0196	0.0172	0.2884	0.0200
18	0.9999	0.0474	0.3105	0.0198	0.0101	0.7274	0.0108
19	0.9679	0.0372	0.2675	0.0142	0.0066	0.5240	0.0070
20	0.9157	0.0340	0.2132	0.0166	0.0046	0.3118	0.0050

also by qualitative means. We also intend to use more diverse and extensive test data to confirm the claims made in this paper.

## References

1. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. *Proc. of ACM SIGMOD (1993)* 207–216
2. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. *Proc. of the 20th VLDB Conference (1994)* 487–499
3. Bayardo, R. J., Agrawal, R.: Mining the Most Interesting Rules. In *Proc. of the 5th ACM SIGKDD (1999)* 145–154
4. Bernadet M.: Basis of a Fuzzy Knowledge Discovery System. In *Proc. of the 4th European Conference on PKDD (2000)* 24–33
5. Clark, P., Boswell, P.: Rule Induction with CN2: Some Recent Improvements. In *Machine Learning: Proc. of the Fifth European Conference (1991)* 151–163
6. Gray, B., Orłowska, M.E.: Ccaia: clustering categorical attributes into interesting association rules. In *Proc. of the 2th Pacific-Asia Conf. on Knowledge Discovery and Data Mining (1998)* 132–143
7. Gyenesei, A.: Mining Weighted Association Rules for Fuzzy Quantitative Items. In *Proc. of the 4th European Conference on PKDD (2000)* 416–423
8. Gyenesei, A.: Determining Fuzzy Sets for Quantitative Attributes in Data Mining Problems. *Proc. of Advances in Fuzzy Systems and Evol. Comp. (2001)* 48–53
9. Hong, T-P., Kuo, C-S, Chi, S-C.: Mining association rules from quantitative data. *Intelligent Data Analysis 3 (5) (1999)* 363–376

**Table 3.** Interestingness measures - ranks

Rule	<i>Conf</i>	<i>Cov</i>	<i>Corr</i>	<i>I-measure</i>	<i>UE-measure</i>	<i>CE-measure</i>	<i>J-measure</i>
1	7	6	5	8	5	6	5
2	10	7	7	9	7	10	7
3	6	4	4	7	4	5	4
4	5	2	2	2	2	4	2
5	13	14	16	14	14	13	14
6	16	11	13	12	11	16	11
7	11	5	6	5	6	11	6
8	9	3	3	3	3	9	3
9	14	15	18	13	16	14	16
10	19	19	19	18	17	19	18
11	4	1	1	1	1	3	1
12	15	9	12	10	9	15	10
13	12	12	14	11	13	12	12
14	18	8	8	4	10	18	8
15	20	20	20	20	20	20	20
16	17	16	17	6	18	17	15
17	8	10	10	16	8	8	9
18	1	13	9	15	12	1	13
19	2	17	11	19	15	2	17
20	3	18	15	17	19	7	19

**Table 4.** Correlation coefficients for interestingness measures

	<i>Conf</i>	<i>Cov</i>	<i>Corr</i>	<i>I-measure</i>	<i>UE-measure</i>	<i>CE-measure</i>	<i>J-measure</i>
<i>Conf</i>	-	0.378	0.500	0.243	0.359	0.949	0.300
<i>Cov</i>	0.378	-	0.987	0.945	0.975	0.372	0.967
<i>Corr</i>	0.500	0.987	-	0.915	0.957	0.495	0.940
<i>I-measure</i>	0.243	0.945	0.915	-	0.919	0.245	0.960
<i>UE-measure</i>	0.359	0.975	0.957	0.919	-	0.380	0.988
<i>CE-measure</i>	0.949	0.372	0.495	0.245	0.380	-	0.326
<i>J-measure</i>	0.300	0.967	0.940	0.960	0.988	0.326	-

10. Hilderman, R.J., Hamilton, H.J.: Knowledge discovery and interestingness measures: A survey. Technical Report CS 99-04, University of Regina, Canada (1999)
11. Kullback, S., Leibler, R.A.: On information and sufficiency. *Annals of Mathematical Statistics*, 22, (1951) 79–86
12. Kuok, C.M., Fu, A., Wong, M.H.: Fuzzy association rules in databases. In *ACM SIGMOD Record* 27(1),(1998) 41–46
13. Morishita, S.: On Classification and Regression. In *Proc. of the First Int. Conf. on Discovery Science - LNAI 1532* (1998) 40–57
14. Piatetsky-Shapiro, G., Frawley, W.J.: *Knowledge Discovery in Databases*. Chapter 13. AAAI Press/The MIT Press, Menlo Park, California (1991)
15. Shannon, C.E., Weaver, W.: *Introduction to Probability and Statistics for Scientists and Engineers*. McGraw-Hill (1997)

16. Smyth, P., Goodman, R.M.: Rule induction using information theory. In Knowledge Discovery in Databases, AAAI/MIT Press (1991) 159–176
17. Srikant, R., Agrawal, R.: Mining quantitative association rules in large relation tables. Proc. of ACM SIGMOD (1996) 1–12
18. Theil, H.: Economics and information theory. North-Holland (1967)