Research Article

Interest Measures for Fuzzy Association Rules Based on Expectations of Independence

Michal Burda

Institute for Research and Applications of Fuzzy Modeling, Centre of Excellence IT4Innovations, University of Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic

Correspondence should be addressed to Michal Burda; michal.burda@osu.cz

Lift, leverage, and conviction are three of the best commonly known interest measures for crisp association rules. All of them are based on a comparison of observed support and the support that is expected if the antecedent and consequent part of the rule were stochastically independent. The aim of this paper is to provide a correct definition of lift, leverage, and conviction measures for fuzzy association rules and to study some of their interesting mathematical properties.

1. Introduction

Searching for association rules is a broadly discussed, developed, and accepted data mining technique [1, 2]. An association rule is an expression $X \rightharpoonup Y$, where antecedent X and consequent Y are conditions, the former usually in the form of elementary conjunction and the latter being usually atomic. Such rules are usually interpreted as the following implicational statement: "if X is satisfied then Y is true very often too."

Naturally, analysts are interested only in such rules that are somehow interesting, unusual, or exceptional. To assess rule interestingness objectively, there have been developed many measures of rule interest or intensity. Among the most essential, *support* and *confidence* are traditionally considered. An objective of association rules mining is to find rules with *support* and *confidence* above some user-defined thresholds.

Searching for association rules fits particularly well on binary or categorical data and many have been written on that topic [1–4]. For association analysis on numeric data, a prior discretization is proposed, for example, by Srikant and Agrawal [5]. Another alternative is to take an advantage of fuzzy logic [6].

The use of fuzzy logic in connection with association rules has been motivated by many authors (see e.g., [7] for recent overview). Fuzzy association rules are appealing also because of the use of vague linguistic terms such as "small" and "very big" [8–11].

In this paper, we focus on three measures of rule intensity that are all based on comparison between the observed support and the support that is expected under the assumption of independence of the rule's antecedent and consequent. These measures are *lift*, *leverage*, and *conviction*. All of them were initially developed for nonfuzzy (i.e., "crisp") association rules.

Lift was firstly described in [12] under its original name "interest." It was well studied for association rules on binary data in [13, 14]. Lift is defined as a ratio of observed support $X \wedge Y$ to the support that is expected under the assumption of independence of X and Y.

On the other hand, *leverage* [15] measures the difference of observed and expected supports. (Hence, it is very similar to lift.)

Conviction [12] is defined as a ratio of expected and observed support of $X \wedge \neg Y$. Although it is very similar to lift, its properties are more similar to confidence. Lallich et al. [16] provide a nice overview of many other crisp rule measures.

As quite many have been written about these measures for crisp association rules, not so much has been done with respect to fuzzy rules. Unfortunately, simplicity of definitions for crisp rules sometimes led to oversimplified definitions for fuzzy rules. Some authors believe the generalization of

those crisp measures for fuzzy data is as trivial as substituting crisp terms with analogous fuzzy terminology inside of crisp-case definitions; see, for example, [17, 18]. Unfortunately, as discussed in this paper, such oversimplification may lead to erroneous outputs. In order to preserve some nice mathematical properties, one must take care of the type of the t-norms being used.

In Section 2, a brief theoretical background for both binary and fuzzy association rules is provided. Section 2.3 discusses shortly naive definitions of lift, leverage, and conviction and shows a simple example where the rational interpretations of these measures become broken. Before providing corrected definitions in Section 4, an essential notion of *expected support* is analysed in Section 3. Finally, Section 5 concludes the paper with summarization of the achieved goals and drawings of the possible directions of future research.

2. Theoretical Background

2.1. Binary Association Rules. Let $\mathcal{O} := \{o_1, o_2, \ldots, o_n\}$, $n > 0$, be a finite set of objects and let $\mathcal{A} := \{a_1, a_2, \ldots, a_m\}$, $m > 0$, be a finite set of attributes (features). Each attribute can be considered as a logical predicate: $a_i(o_j)$ is true (or false) accordingly to whether the ith attribute applies (or does not apply) to object o_j. For a subset $X \subseteq \mathcal{A}$ of attributes, let us define a new predicate of a logical conjunction of the attributes contained in X as follows:

$$X\left(o_j\right) :\equiv \forall a_i \in X : a_i\left(o_j\right). \tag{1}$$

Moreover, let us define a negated predicate $\neg X(o_j)$ as follows:

$$\neg X\left(o_j\right) :\equiv \exists a_i \in X : \neg a_i\left(o_j\right). \tag{2}$$

An association rule is a formula $X \rightharpoonup Y$, where $X \subset \mathcal{A}$ is an *antecedent*, $Y \subset \mathcal{A}$ is a *consequent,* and $X \cap Y = \emptyset$. (Typically, $|X| \geq 1$ and $|Y| = 1$.) Both X and Y are sometimes called *itemsets*. Please consider the following rule as an example:

$$\{\text{tequila, salt}\} \longrightarrow \{\text{lemon}\}. \tag{3}$$

The *support* and *confidence* are defined as follows [2, 19]:

$$\text{supp}\left(X\right) := \frac{|\{o \in \mathcal{O} \mid X\left(o\right)\}|}{n}, \tag{4}$$

$$\text{supp}\left(X \rightharpoonup Y\right) := \frac{|\{o \in \mathcal{O} \mid X\left(o\right) \wedge Y\left(o\right)\}|}{n}, \tag{5}$$

$$\text{conf}\left(X \rightharpoonup Y\right) := \frac{\text{supp}\left(X \rightharpoonup Y\right)}{\text{supp}\left(X\right)}, \tag{6}$$

where $n = |\mathcal{O}|$. Evidently, $\text{supp}(\neg X) = 1 - \text{supp}(X)$.

If $\mathcal{O}$ is a random sample, then observing $o \in \mathcal{O}$ is a random event. Then also a random event X may be defined on the basis of the truth value of the predicate $X(o)$ and support $\text{supp}(X)$ becomes an estimate of a probability $P(\mathsf{X})$. Then also $\text{supp}(\neg X)$ would correspond to the probability $P(\overline{\mathsf{X}})$, where

$\overline{\mathsf{X}}$ is complementary event to X. Confidence $\text{conf}(X \rightharpoonup Y)$ would then be an estimate of conditional probability $P(\mathsf{Y} \mid \mathsf{X})$.

Lift [12], *leverage* [15], and *conviction* [12] for binary data are defined as follows:

$$\text{lift}\left(X \rightharpoonup Y\right) := \frac{\text{supp}\left(X \rightharpoonup Y\right)}{\text{supp}\left(X\right) \cdot \text{supp}\left(Y\right)} = \frac{\text{conf}\left(X \rightharpoonup Y\right)}{\text{supp}\left(Y\right)}; \tag{7}$$

$$\text{lever}\left(X \rightharpoonup Y\right) := \text{supp}\left(X \rightharpoonup Y\right) - \text{supp}\left(X\right) \cdot \text{supp}\left(Y\right); \tag{8}$$

$$\text{conv}\left(X \rightharpoonup Y\right) := \frac{\text{supp}\left(X\right) \cdot \text{supp}\left(\neg Y\right)}{\text{supp}\left(X \rightharpoonup \neg Y\right)} = \frac{1 - \text{supp}\left(Y\right)}{1 - \text{conf}\left(X \rightharpoonup Y\right)}. \tag{9}$$

If X and Y are stochastically independent, $P(\mathsf{X} \wedge \mathsf{Y}) = P(\mathsf{X}) \cdot P(\mathsf{Y})$; that is, the expression $\text{supp}(X) \cdot \text{supp}(Y)$ is an estimation of support $\text{supp}(X \rightharpoonup Y)$ under the assumption of X and Y being independent.

Hence, lift $\text{lift}(X \rightharpoonup Y)$ is a ratio of observed support to the support that is expected under the assumption of independence, leverage $\text{lever}(X \rightharpoonup Y)$ is a difference of observed and expected support, and conviction is a ratio of the expected support of X appearing without Y to the observed support $\text{supp}(X \rightharpoonup \neg Y)$.

If X and Y are independent, $\text{lift}(X \rightharpoonup Y) \simeq 1$, $\text{conv}(X \rightharpoonup Y) \simeq 1$, and $\text{lever}(X \rightharpoonup Y) \simeq 0$. The values of $\text{lift}(X \rightharpoonup Y) > 1$, $\text{conv}(X \rightharpoonup Y) > 1$, and $\text{lever}(X \rightharpoonup Y) > 0$ indicate positive relationship, while $\text{lift}(X \rightharpoonup Y) < 1$, $\text{conv}(X \rightharpoonup Y) < 1$, and $\text{lever}(X \rightharpoonup Y) < 0$ indicate negative relationship.

2.2. Fuzzy Association Rules. For fuzzy association rules, domain of each fuzzy attribute $a \in \mathcal{A}$ is not binary (or "crisp") $\{0, 1\}$ but graded (or "fuzzy"), that is, interval $[0, 1]$. That is, for each $a \in \mathcal{A}$ and $o \in \mathcal{O}$, $a(o) \in [0, 1]$. For a subset $X \subseteq \mathcal{A}$ of fuzzy attributes, we define a new predicate of a logical conjunction (similarly to binary case (1)) by using a t-norm $\otimes$ as

$$X\left(o_j\right) := \bigotimes_{a \in X} a\left(o_j\right). \tag{10}$$

T-norm $\otimes$ is a generalized logical conjunction, that is, a function $[0, 1] \times [0, 1] \rightarrow [0, 1]$ which is associative, commutative, and monotone increasing (in both places) and which satisfies the boundary conditions $\alpha \otimes 0 = 0$ and $\alpha \otimes 1 = \alpha$ for each $\alpha \in [0, 1]$. Some well-known examples of t-norms are as follows:

(i) product t-norm: $\otimes_{\text{prod}}(\alpha, \beta) = \alpha\beta$;

(ii) minimum t-norm: $\otimes_{\text{min}}(\alpha, \beta) = \min(\alpha, \beta)$;

(iii) Łukasiewicz t-norm: $\otimes_{\text{Łuk}}(\alpha, \beta) = \max(0, \alpha + \beta - 1)$.

Negation of fuzzy predicate a is often defined for any $o \in \mathcal{O}$ as $\neg a(o) := 1 - a(o)$. Then, we can define (similarly to (2)) a negated predicate $\neg X$ as follows:

$$\neg X\left(o_j\right) := 1 - X\left(o_j\right). \tag{11}$$

Let $a \in \mathscr{A}$, $o \in \mathscr{O}$, $n = |\mathscr{O}|$, $n > 0$, $X, Y \subset \mathscr{A}$, $X \neq \emptyset$, $Y \neq \emptyset$, and $X \cap Y = \emptyset$. Several intensity measures may be defined as follows:

$$\text{fsupp}(X) := \frac{\sum_{o \in \mathscr{O}} X(o)}{n},$$

$$\text{fsupp}(X \rightharpoonup Y) := \frac{\sum_{o \in \mathscr{O}} (X(o) \otimes Y(o))}{n}, \qquad (12)$$

$$\text{fconf}(X \rightharpoonup Y) := \frac{\text{fsupp}(X \rightharpoonup Y)}{\text{fsupp}(X)}.$$

Convention 1. Throughout this text, we assume $\otimes$ is an arbitrary (but fixed) t-norm. Where it is important to explicitly specify a concrete t-norm, say $\otimes_{\text{prod}}$, we put $\otimes_{\text{prod}}$ in subscript and write, for example, $\text{fsupp}_{\otimes_{\text{prod}}}(X \rightharpoonup Y)$ instead of $\text{fsupp}(X \rightharpoonup Y)$ or $X_{\otimes_{\text{prod}}}(o)$ instead of $X(o)$.

Convention 2. For the sake of simplicity, we will sometimes express a fuzzy attribute as a vector of membership degrees. For instance, suppose $a \in \mathscr{A}$ is a fuzzy attribute on a set of objects $\mathscr{O} = \{o_1, o_2, o_3\}$. Then, instead of writing $a(o_1) := 0.3$, $a(o_2) := 0.9$, and $a(o_3) := 1$, we will use a much concise form: $a := (0.3, 0.9, 1)$.

Example 1. Let $a := (a_1, a_2, a_3)$, $b := (b_1, b_2, b_3)$, and $c := (c_1, c_2, c_3)$ be fuzzy attributes. Then,

$$\text{fsupp}_{\otimes_{\text{prod}}}(\{a, b\} \rightharpoonup \{c\}) = \frac{1}{3} \sum_{i=1}^{3} \left(a_i \otimes_{\text{prod}} b_i \otimes_{\text{prod}} c_i \right). \qquad (13)$$

2.3. Naive Definition of Lift, Leverage, and Conviction for Fuzzy Association Rules. A naive approach for introducing lift, leverage, and conviction into the fuzzy association rules framework is to use simply their definitions (7), (8), and (9) for binary rules and replace binary support (4) and (5) and confidence (6) with their fuzzy alternatives (12) as, for example, in [17, 18]. Unfortunately, that approach works well only for $\otimes$ being the product t-norm. As indicated in the following experiment, using minimum or Łukasiewicz t-norms may lead to erroneous interpretations.

Experiment 1. Two vectors X and Y of size $n = 2000$ were randomly generated from the uniform distribution on the interval $[0, 1]$ so that they are stochastically independent. Above-described naive versions of lift, leverage, and conviction of a rule $X \rightharpoonup Y$ were computed with using minimum, product, and Łukasiewicz t-norms as $\otimes$; see Table 1 for results.

As discussed in Section 2.1, stochastically independent data are expected to result in lift and conviction being close to 1 while leverage is expected to be close to 0. As can be seen, this is the case only for product t-norm.

The values of naive lift and naive leverage wrongly indicate positive (resp., negative) relationship, if the minimum (resp., Łukasiewicz) t-norm is used. Paradoxically, naive conviction indicates opposite sign of relationship.

Table 1: Comparison of lift, leverage, and conviction computed with different t-norms on stochastically independent data generated randomly from uniform distribution.

	Naive lift	Naive leverage	Naive conviction
Łukasiewicz t-norm	0.675	−0.081	1.544
Product t-norm	1.012	0.003	1.012
Minimum t-norm	1.353	0.088	0.755

3. Expected Support of a Conjunction of Fuzzy Attributes under the Assumption of Independence

As indicated in Experiment 1 presented in Section 2.3 above, naive lift, leverage, and conviction no more behave like their alternatives for crisp data: they no more represent a ratio of what is observed to what is expected under the assumption of independence. To recover their definitions, independency of fuzzy attributes must be treated correctly. Only then a proper definitions of lift, leverage, and conviction can be formulated.

Given sets X and Y of fuzzy attributes, what support of $X \wedge Y$ is expected if X and Y are independent? Moreover, what does independency of fuzzy attributes mean?

For the sake of simplicity, let us assume X and Y are sets containing a single fuzzy attribute; that is, $|X| = |Y| = 1$. For more complex cases, a new attribute can be created from the set of fuzzy attributes by using (10).

If $\mathscr{O}$ is a set of randomly selected objects, one can consider membership values $X(o)$ and $Y(o)$ as random variables X and Y and treat the independence of fuzzy attributes as stochastic independence of random variables X and Y.

Two random variables X, Y are stochastically independent, if the combined random variable (X, Y) has a joint probability density as

$$f_{\mathsf{X},\mathsf{Y}}(x, y) = f_{\mathsf{X}}(x) f_{\mathsf{Y}}(y). \qquad (14)$$

If X and Y are two independent random variables from interval $[0, 1]$, then

$$\sigma(x, y) := \frac{x \otimes y}{n} \qquad (15)$$

is a random variable with probability density function $f_{\sigma}(x, y) = f_{\mathsf{X},\mathsf{Y}}(x, y)$.

By definition, *expected value* $E[\mathsf{Z}]$ of a random variable Z is a weighted average of all possible values. More formally,

$$E[\mathsf{Z}] = \int_{-\infty}^{\infty} z f_{\mathsf{Z}}(z) \, \mathrm{d}z, \qquad (16)$$

where f_{Z} is a probability density function of random variable Z.

Similarly, an expected value $E[\sigma(x, y)]$ is a weighted average of all possible (x, y) pairs, namely,

$$E[\sigma(x, y)] = \iint_0^1 \sigma(x, y) f_{\sigma}(x, y) \, \mathrm{d}x \, \mathrm{d}y. \qquad (17)$$

In real setting, $f_{\sigma}(x, y)$ is unknown but we can estimate its values from data (i.e., from objects $\mathscr{O}$ and their fuzzy

attributes $\mathscr{A}$) by using the assumption of independence (14) as follows:

$$f_\sigma(x, y) = f_X(x) f_Y(y) \approx \frac{\text{count}_X(x)}{n} \cdot \frac{\text{count}_Y(y)}{n}, \quad (18)$$

where $\text{count}_A(a)$ is the number of objects from $\mathscr{O}$ that belong to $A \in \mathscr{A}$ with degree a; that is, $\text{count}_A(a) = |\{o \in \mathscr{O} \mid A(o) = a\}|$.

Assuming $x \in \{X(o) \mid o \in \mathscr{O}\}$ and $y \in \{Y(o) \mid o \in \mathscr{O}\}$, we obtain, from (15), (17), and (18),

$$E[\sigma(x, y)] \approx \iint_0^1 \frac{(x \otimes y) \cdot \text{count}_X(x) \cdot \text{count}_Y(y)}{n^3} dx\, dy$$

$$= \sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \otimes Y(o_j)}{n^3}. \tag{19}$$

Since X and Y are independent, $E[\text{fsupp}(X \rightharpoonup Y)] = n \cdot E[\sigma(x, y)]$ and hence expected value of $\text{fsupp}(X \rightharpoonup Y)$ is

$$E[\text{fsupp}(X \rightharpoonup Y)] \approx \sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \otimes Y(o_j)}{n^2}. \tag{20}$$

Now, we are ready to define notions of *expected support* and *expected confidence*.

Definition 2. Let $\otimes$ be a t-norm, and let X, Y be sets of fuzzy attributes such that $\text{fsupp}(X) > 0$ and $n > 0$. Then, the *expected fuzzy support* $\widehat{\text{fsupp}}(X \rightharpoonup Y)$ and the *expected fuzzy confidence* $\widehat{\text{fconf}}(X \rightharpoonup Y)$ are defined as follows:

$$\widehat{\text{fsupp}}(X \rightharpoonup Y) := \sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \otimes Y(o_j)}{n^2},$$

$$\widehat{\text{fconf}}(X \rightharpoonup Y) := \frac{\widehat{\text{fsupp}}(X \rightharpoonup Y)}{\text{fsupp}(X)}. \tag{21}$$

Proposition 3. *Let X, Y be sets of fuzzy attributes. Then,*

(1) *if $fsupp(X \rightharpoonup Y) > 0$, then $\widehat{fsupp}(X \rightharpoonup Y) > 0$,*

(2) *$\widehat{fsupp}(X \rightharpoonup Y) \leq \min(fsupp(X), fsupp(Y))$.*

Proof. (1) If $\text{fsupp}(X \rightharpoonup Y) > 0$, then

$$\sum_{i=1}^n X(o_i) \otimes Y(o_i) > 0 \tag{22}$$

and hence

$$\sum_{i=1}^n \sum_{j=1}^n X(o_i) \otimes Y(o_j) > 0. \tag{23}$$

Therefore also $\widehat{\text{fsupp}}(X \rightharpoonup Y) > 0$.

(2) For any t-norm $\otimes$ holds that $X(o_i) \otimes Y(o_j) \leq Y(o_j)$. Therefore,

$$\sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \otimes Y(o_j)}{n^2} \leq \sum_{i=1}^n \sum_{j=1}^n \frac{Y(o_j)}{n^2} = \text{fsupp}(Y) \tag{24}$$

and similarly for $\text{fsupp}(X)$. Hence, $\widehat{\text{fsupp}}(X \rightharpoonup Y) \leq \min(\text{fsupp}(X), \text{fsupp}(Y))$. $\square$

Remark 4. Note that the reverse direction of the first implication of Proposition 3 is not generally true. For example, for $X := (1, 0)$ and $Y := (0, 1)$, we have $\widehat{\text{fsupp}}(X \rightharpoonup Y) = 1/4$ but $\text{fsupp}(X \rightharpoonup Y) = 0$.

Proposition 5. *Let X, Y be sets of fuzzy attributes and let $\otimes := \otimes_{prod}$; that is, the product t-norm is being used as a conjunction. Then,*

$$\widehat{fsupp}(X \rightharpoonup Y) = fsupp(X) \cdot fsupp(Y). \tag{25}$$

Proof. If $\otimes := \otimes_{\text{prod}}$, $\widehat{\text{fsupp}}(X \rightharpoonup Y)$ equals

$$\sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \cdot Y(o_j)}{n^2}$$

$$= \frac{1}{n^2} \sum_{i=1}^n X(o_i) \sum_{i=1}^n Y(o_i) = \text{fsupp}(X) \cdot \text{fsupp}(Y). \tag{26}$$

$\square$

Proposition 6. *Let X, Y be sets of fuzzy attributes and let $\otimes := \otimes_{min}$; that is, the minimum t-norm is being used as a conjunction. Then,*

$$fsupp(X) \cdot fsupp(X) \leq \widehat{fsupp}(X \rightharpoonup Y)$$

$$\leq fsupp(X) \otimes_{min} fsupp(Y). \tag{27}$$

Proof. Let $\otimes := \otimes_{\min}$. The first inequality follows from the fact that $\otimes_{\min}(x, y) \geq x \cdot y$ as

$$\widehat{\text{fsupp}}(X \rightharpoonup Y) = \sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \otimes_{\min} Y(o_j)}{n^2}$$

$$\geq \sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \cdot Y(o_j)}{n^2} \tag{28}$$

$$= \text{fsupp}(X) \cdot \text{fsupp}(Y).$$

The second inequality follows directly from Proposition 3. $\square$

Proposition 7. *Let X, Y be sets of fuzzy attributes and let $\otimes := \otimes_{Łuk}$; that is, the Łukasiewicz t-norm is being used as a conjunction. Then,*

$$fsupp(X) \otimes_{Łuk} fsupp(Y) \leq \widehat{fsupp}(X \rightharpoonup Y)$$

$$\leq fsupp(X) \cdot fsupp(Y). \tag{29}$$

Proof. Let $\otimes := \otimes_{Łuk}$, $s_X := \text{fsupp}(X)$, and $s_Y := \text{fsupp}(Y)$. To prove the first inequality, it suffices to prove

$$\widehat{\text{fsupp}}(X \rightharpoonup Y) \geq \max(0, s_X + s_Y - 1) \tag{30}$$

which is obvious for $s_X + s_Y \leq 1$. Let us therefore assume $s_X + s_Y > 1$; then, (30) can be rewritten as

$$\sum_{i=1}^n \sum_{j=1}^n \frac{X(o_i) \otimes_{Łuk} Y(o_j)}{n^2} \geq s_X + s_Y - 1. \tag{31}$$

Then,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} X\left(o_i\right) \otimes_{\text{Łuk}} Y\left(o_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(X\left(o_i\right) + Y\left(o_j\right)\right) - \sum_{\forall k}\left(t_k\right) - |r|, \tag{32}$$

where t is a sequence of numbers $(X(o_i) + Y(o_j) \mid X(o_i) + Y(o_j) < 1)$ and $r = \{(i, j) \mid X(o_i) + Y(o_j) \geq 1\}$, for $i, j \in \{1, 2, \ldots, n\}$. Since $|t| + |r| = n^2$ and each $t_k < 1$, we can immediately see that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} X\left(o_i\right) \otimes_{\text{Łuk}} Y\left(o_j\right)$$

$$\geq \sum_{i=1}^{n}\sum_{j=1}^{n}\left(X\left(o_i\right) + Y\left(o_j\right)\right) - n^2 \tag{33}$$

$$= n^2 s_X + n^2 s_Y - n^2 = n^2\left(s_X + s_Y - 1\right);$$

hence, (30) holds.

The second inequality follows from $\otimes_{\text{Łuk}}(x, y) \leq x \cdot y$ because then we have

$$\widehat{\text{fsupp}}\left(X \rightharpoonup Y\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X\left(o_i\right) \otimes_{\text{Łuk}} Y\left(o_j\right)}{n^2}$$

$$\leq \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X\left(o_i\right) \cdot Y\left(o_j\right)}{n^2} = s_X \cdot s_Y. \tag{34}$$

$$\square$$

4. Correct Definition of Lift, Leverage, and Conviction for Fuzzy Association Rules

Now, we are ready to provide a correct definition of lift, leverage, and conviction in the framework of fuzzy association rules.

Definition 8. Let X, Y be sets of fuzzy attributes with $n := |\mathcal{O}| > 0$. Then, lift, leverage, and conviction of a fuzzy association rule $X \rightharpoonup Y$ are defined as follows:

$$\text{flift}\left(X \rightharpoonup Y\right) := \frac{\text{fsupp}\left(X \rightharpoonup Y\right)}{\widehat{\text{fsupp}}\left(X \rightharpoonup Y\right)};$$

$$\text{flever}\left(X \rightharpoonup Y\right) := \text{fsupp}\left(X \rightharpoonup Y\right) - \widehat{\text{fsupp}}\left(X \rightharpoonup Y\right); \tag{35}$$

$$\text{fconv}\left(X \rightharpoonup Y\right) := \frac{\widehat{\text{fsupp}}\left(X \rightharpoonup \neg Y\right)}{\text{fsupp}\left(X \rightharpoonup \neg Y\right)}.$$

Let us now study some interesting properties of the newly defined notions.

Proposition 9. *Let X, Y be sets of fuzzy attributes. Then,*

(1) $flift(X \rightharpoonup Y) = flift(Y \rightharpoonup X);$

(2) $flift(X \rightharpoonup Y) = fconf(X \rightharpoonup Y)/\widehat{fconf}(X \rightharpoonup Y);$

(3) $0 \leq flift(X \rightharpoonup Y) \leq n;$

(4) *if $fsupp(X \rightharpoonup Y) > 0$, then $flift(X \rightharpoonup Y) > 0$.*

Proof. (1) and (2) directly follow from the definitions and from the fact that t-norms are commutative.

(3) Since the membership degrees are defined on interval $[0, 1]$, their sums cannot be negative either. Hence, flift$(X \rightharpoonup Y) \geq 0$. Next, assume to the contrary that flift$(X \rightharpoonup Y) > n$ as follows:

$$\text{flift}\left(X \rightharpoonup Y\right) = n\frac{\sum_{i=1}^{n} X\left(o_i\right) \otimes Y\left(o_i\right)}{\sum_{i=1}^{n}\sum_{j=1}^{n} X\left(o_i\right) \otimes Y\left(o_j\right)} > n; \tag{36}$$

therefore,

$$\sum_{i=1}^{n} X\left(o_i\right) \otimes Y\left(o_i\right) > \sum_{i=1}^{n}\sum_{j=1}^{n} X\left(o_i\right) \otimes Y\left(o_j\right), \tag{37}$$

which is a contradiction.

(4) If fsupp$(X \rightharpoonup Y) > 0$, then, from Proposition 3, we know that also $\widehat{\text{fsupp}}(X \rightharpoonup Y) > 0$. Therefore, flift$(X \rightharpoonup Y)$ exists and is greater than 0. $\square$

Proposition 10. *Let X, Y be sets of fuzzy attributes. Then,*

(1) $flever(X \rightharpoonup Y) = flever(Y \rightharpoonup X),$

(2) $flever(X \rightharpoonup Y) = fsupp(X)(fconf(X \rightharpoonup Y) - \widehat{fconf}(X \rightharpoonup Y)),$

(3) $1/n - 1 \leq flever(X \rightharpoonup Y) \leq fsupp(X \rightharpoonup Y)(1 - 1/n).$

Proof. (1) and (2) directly follow from the definitions and from the fact that t-norms are commutative.

(3) Let $u := \sum_{\forall i} X(o_i) \otimes Y(o_i)$ and $r := \sum_{\forall i}\sum_{\forall j \neq i} X(o_i) \otimes Y(o_j)$. Evidently, fsupp$(X \rightharpoonup Y) = u/n$ and $\widehat{\text{fsupp}}(X \rightharpoonup Y) = (u+r)/n^2$. Then, flever$(X \rightharpoonup Y) = u/n - (u+r)/n^2$. Obviously, $0 \leq u \leq n$ and $0 \leq r \leq n^2 - n$. Therefore,

$$-\frac{n^2 - n}{n^2} = \frac{1}{n} - 1$$

$$\leq \text{flever}\left(X \rightharpoonup Y\right) \leq \frac{u}{n} - \frac{u}{n^2} \tag{38}$$

$$= \text{fsupp}\left(X \rightharpoonup Y\right)\left(1 - \frac{1}{n}\right).$$

$$\square$$

Proposition 11. *Let X, Y be sets of fuzzy attributes. Then,*

(1) $fconv(X \rightharpoonup Y) = 1/flift(X \rightharpoonup \neg Y);$

(2) $fconv(X \rightharpoonup Y) = fconv(\neg Y \rightharpoonup \neg X);$

(3) $fconv(X \rightharpoonup Y) = \widehat{fconf}(X \rightharpoonup \neg Y)/fconf(X \rightharpoonup \neg Y);$

(4) $1/n \leq fconv(X \rightharpoonup Y).$

Proof. Everything directly follows from Definition 8 and Proposition 9. $\square$

Corollary 12. *Let X, Y be sets of fuzzy attributes and let $\otimes :=$ $\otimes_{prod}$; that is, the product t-norm is being used as a conjunction. Then,*

$$flift(X \rightharpoonup Y) = \frac{fsupp(X \rightharpoonup Y)}{fsupp(X) \cdot fsupp(Y)} = \frac{fconf(X \rightharpoonup Y)}{fsupp(Y)};$$

$$flever(X \rightharpoonup Y) = fsupp(X \rightharpoonup Y) - fsupp(X) \cdot fsupp(Y);$$

$$fconv(X \rightharpoonup Y) = \frac{fsupp(X) \cdot fsupp(\neg Y)}{fsupp(X \rightharpoonup \neg Y)}$$

$$= \frac{1 - fsupp(Y)}{1 - fconf(X \rightharpoonup Y)}.$$

$$(39)$$

Proof. Everything directly follows from Definitions 2 and 8, from Proposition 5, and from the fact that $fsupp(\neg Y) = 1 - fsupp(Y)$. □

Corollary 12 copies properties that are well known for crisp variants of lift, leverage, and conviction. In practice, the use of these equations is much more convenient than the original ones from Definition 8. However, note that Corollary 12 holds only if product t-norm $\otimes_{prod}$ is used. For other t-norms such as minimum $\otimes_{min}$ or Łukasiewicz $\otimes_{Łuk}$, Definition 8 must not be oversimplified that way. See the subsequent corollaries for more details.

Corollary 13. *Let X, Y be sets of fuzzy attributes and let $\otimes := \otimes_{min}$; that is, the minimum t-norm is being used as a conjunction. Then,*

$$\frac{s_{XY}}{s_X \otimes_{min} s_Y} \leq flift(X \rightharpoonup Y) \leq \frac{s_{XY}}{s_X \cdot s_Y},$$

$$s_{XY} - s_X \otimes_{min} s_Y \leq flever(X \rightharpoonup Y) \leq s_{XY} - s_X \cdot s_Y, \quad (40)$$

$$\frac{s_X \cdot s_{\overline{Y}}}{s_{X\overline{Y}}} \leq fconv(X \rightharpoonup Y) \leq \frac{s_X \otimes_{min} s_{\overline{Y}}}{s_{X\overline{Y}}},$$

where $s_X := fsupp(X)$, $s_Y := fsupp(Y)$, $s_{\overline{Y}} := fsupp(\neg Y)$, $s_{XY} := fsupp(X \rightharpoonup Y)$, and $s_{X\overline{Y}} := fsupp(X \neg Y)$.

Proof. Everything directly follows from Definitions 2 and 8 and from Propositions 6 and 11. □

Corollary 14. *Let X, Y be sets of fuzzy attributes and let $\otimes := \otimes_{Łuk}$; that is, the Łukasiewicz t-norm is being used as a conjunction. Then,*

$$\frac{s_{XY}}{s_X \cdot s_Y} \leq flift(X \rightharpoonup Y) \leq \frac{s_{XY}}{s_X \otimes_{Łuk} s_Y};$$

$$s_{XY} - s_X \cdot s_Y \leq flever(X \rightharpoonup Y) \leq s_{XY} - s_X \otimes_{Łuk} s_Y; \quad (41)$$

$$\frac{s_X \otimes_{Łuk} s_{\overline{Y}}}{s_{X\overline{Y}}} \leq fconv(X \rightharpoonup Y) \leq \frac{s_X \cdot s_{\overline{Y}}}{s_{X\overline{Y}}},$$

where $s_X := fsupp(X)$, $s_Y := fsupp(Y)$, $s_{\overline{Y}} := fsupp(\neg Y)$, $s_{XY} := fsupp(X \rightharpoonup Y)$, and $s_{X\overline{Y}} := fsupp(X \rightharpoonup \neg Y)$.

TABLE 2: Comparison of lift, leverage, and conviction computed with different t-norms on stochastically independent data generated randomly from uniform distribution.

	Naive lift	Naive leverage	Naive conviction
Łukasiewicz t-norm	1.022	0.004	1.021
Product t-norm	1.012	0.003	1.012
Minimum t-norm	1.010	0.003	1.011

Proof. Everything directly follows from Definitions 2 and 8 and from Propositions 7 and 11. □

Experiment 2. The same data as in Experiment 1 were processed accordingly to correct definitions of lift, leverage, and conviction (see Definition 8). The results can be found in Table 2. Since the data are randomly generated from uniform distribution, they are stochastically independent; hence, lift and conviction should be close to 1 and lift should be close to 0 regardless of t-norm being used. As one can see, the results in Table 2 are perfectly in compliance with our expectations.

5. Conclusion

Lift is a ratio of observed support (resp., confidence) to the support (resp., confidence) that is expected under the assumption of independence. Leverage is similar to lift, since it is a difference between observed and expected support. Conviction is often treated as an alternative to confidence; nevertheless, it is defined on the basis of observed and expected support too.

In this paper, a correct definition of lift, leverage, and conviction for fuzzy data was provided. It should be stressed here that there already exist some research papers that use incorrect (a.k.a. "naive") definition of fuzzy lift (e.g., [18]) that is also discussed here.

The naive definition does not preserve interpretation of positive/negative relationship. For crisp lift and conviction (resp., leverage), the stochastically independent attributes X and Y yield $lift(X \rightharpoonup Y) \approx 1$, $conv(X \rightharpoonup Y) \approx 1$, and $lever(X \rightharpoonup Y) \approx 0$. As shown in Experiment 1, this is no more the case for naive definitions of these measures. On the other hand, Experiment 2 shows that correct definitions of lift, leverage, and conviction again recover that feature for fuzzy data.

All the lift, leverage, and conviction definitions are similar to their "crisp" alternatives (i.e., definitions for binary data) if the t-norm being used is the product $\otimes_{prod}$. For Łukasiewicz $\otimes_{Łuk}$ and minimum $\otimes_{min}$ t-norms, a more complicated computation takes place.

In [20], an algorithm was developed in for fast evaluation of fuzzy lift. A future research will therefore address improvements of association rules search algorithms by introducing fast computations of other measures, adding pruning heuristics based on boundary conditions provided by Corollaries 13 and 14. Also, other interest measures may be studied and their applicability to fuzzy rules may be considered.

Conflict of Interests

Acknowledgment

References

[1] M. Hahsler and K. Hornik, "New probabilistic interest measures for association rules," *Intelligent Data Analysis*, vol. 11, no. 5, pp. 437–455, 2007.

[2] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large databases," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 207–216, Washington, DC, USA, May 1993.

[3] A. Berrado and G. C. Runger, "Using metarules to organize and group discovered association rules," *Data Mining and Knowledge Discovery*, vol. 14, no. 3, pp. 409–431, 2007.

[4] G. I. Webb, "Discovering significant patterns," *Machine Learning*, vol. 68, no. 1, pp. 1–33, 2007.

[5] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," *SIGMOD Record*, vol. 25, no. 2, pp. 1–12, 1996.

[6] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Mass, USA, 1999.

[7] H. Kalia, S. Dehuri, and A. Ghosh, "A survey on fuzzy association rule mining," *International Journal of Data Warehousing and Mining*, vol. 9, no. 1, pp. 1–27, 2013.

[8] K. C. C. Chen and W.-H. Au, *Mining Fuzzy Association Rules*, ACM, New York, NY, USA, 1997.

[9] V. Novák, I. Perfilieva, A. Dvořák, G. Chen, Q. Wei, and P. Yan, "Mining pure linguistic associations from numerical data," *International Journal of Approximate Reasoning*, vol. 48, no. 1, pp. 4–22, 2008.

[10] M. Burda, "Fast evaluation of t-norms for fuzzy association rules mining," in *Proceedings of the 14th IEEE International Symposium on Computational Intelligence and Informatics (CINTI '13)*, pp. 465–470, Budapest, Hungary, November 2013.

[11] M. Burda, V. Pavliska, and R. Valášek, "Parallel mining of fuzzy association rules on dense data sets," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '14)*, pp. 2156–2162, IEEE, Beijing, China, 2014.

[12] S. Brin, R. Motwani, J. D. Ullman, and S. Tsur, "Dynamic itemset counting and implication rules for market basket data," in *Proceedings ACM SIGMOD International Conference on Management of Data (SIGMOD '97)*, pp. 255–264, Tucson, Ariz, USA, 1997.

[13] P. D. McNicholas, T. B. Murphy, and M. O'Regan, "Standardising the lift of an association rule," *Computational Statistics & Data Analysis*, vol. 52, no. 10, pp. 4712–4721, 2008.

[14] P. Hájek, I. Havel, and M. Chytil, "The GUHA method of automatic hypotheses determination," *Computing*, vol. 1, no. 4, pp. 293–308, 1966.

[15] G. Piatetsky-Shapiro, "Discovery, analysis, and presentation of strong rules," in *Knowledge Discovery in Databases*, pp. 229–248, AAAI/MIT Press, 1991.

[16] S. Lallich, O. Teytaud, and E. Prudhomme, "Association rule interestingness: measure and statistical validation," in *Quality Measures in Data Mining*, vol. 43 of *Studies in Computational Intelligence*, pp. 251–275, Springer, Berlin, Germany, 2007.

[17] A. L. Buczak and C. M. Gifford, "Fuzzy association rule mining for community crime pattern discovery," in *Proceedings of the ACM SIGKDD Workshop on Intelligence and Security Informatics (ISI-KDD '10)*, vol. 2, pp. 1–10, ACM, New York, NY, USA, 2010.

[18] D. P. Pancho, J. M. Alonso, J. Alcalá-Fdez, and L. Magdalena, "Interpretability analysis of fuzzy association rules supported by fingrams," in *Proceedings of the 8th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT '13)*, pp. 469–474, September 2013.

[19] R. Agrawal, "Fast discovery of association rules," in *Advances in Knowledge Discovery and Data Mining*, pp. 307–328, AAAI Press/MIT Press, Cambridge, Mass, USA, 1996.

[20] M. Burda, "Lift measure for fuzzy association rules," in *Proceedings of the 7th International Conference on Soft Methods in Probability and Statistics (SMPS '2014)*, Springer, Warsaw, Poland, September 2014.