

Article

Data Mining Using Association Rules for Intuitionistic Fuzzy Data

Frederick Petry ^{1,*}  and Ronald Yager ²
¹ Cognitive Geospatial Systems, Naval Research Laboratory, Stennis Space Center, MS 39529, USA

² Machine Intelligence Institute, Iona College, New Rochelle, NY 10804, USA; yager@panix.com

* Correspondence: frednavy3@gmail.com

Abstract: This paper considers approaches to the computation of association rules for intuitionistic fuzzy data. Association rules can provide guidance for assessing the significant relationships that can be determined while analyzing data. The approach uses the cardinality of intuitionistic fuzzy sets that provide a minimum and maximum range for the support and confidence metrics. A new notation is used to enable the representation of the fuzzy metrics. A running example of queries about the desirable features of vacation locations is used to illustrate.

Keywords: data mining; association rules; support metric; confidence metric; Apriori algorithm; intuitionistic fuzzy sets

1. Introduction

Knowledge discovery and data mining involve a number of approaches from the areas of database processing, pattern recognition, and machine learning. These techniques are used in order to find data patterns or associations of value for decision-making. To deal with the uncertainty often found in real-world data, there have been significant considerations about data mining involving uncertain and fuzzy data mining [1–4]. The use of fuzzy sets is well established in managing real-world applications in which uncertainty is commonly involved. More specifically for our interest here in association rules, a number of approaches have been considered for fuzzy data [5–7]. While fuzzy sets present some capabilities, a simple membership function has limitations [8]. This motivates our use of intuitionistic fuzzy sets, as they provide more flexibility for human interactions under uncertainty in association rule data mining. Some concepts are more easily approached by separately envisaging positive and negative instances. In preference modeling involving various interrelated factors, it can be very difficult to simply specify a single simple membership function. [9] For example, in formulating preferences for a family vacation location, there are usually multiple criteria such as distance, costs, and location desirability which influence the evaluation. It can then be very difficult to formulate a simple degree of suitability, but by specifying membership and non-membership corresponding to preferences, the choices can be more easily modeled.

This paper considers the modifications required for association rule computations if fuzzy intuitionistic valued information is involved. The goal is to determine useful associations and patterns from large data sources. Association rules are used to develop valuable insights in determining significant correlations that can be found in the current environment of large data sets, such as those found in various databases or the cloud [10]. The extension using intuitionistic sets is very significant as it captures both positive and negative evaluations. This can then provide contrasting association rules to better inform decision-making. Additionally this provides a complementary capability to the approaches in which positive and negative association rules are generated [11,12].

There have been a large number of specific data mining and knowledge discovery algorithms that have been designed and implemented [13]. Patterns and relationships that can be discovered must be assessed based on interestingness measures in order to provide



Citation: Petry, F.; Yager, R. Data Mining Using Association Rules for Intuitionistic Fuzzy Data. *Information* **2023**, *14*, 372. <https://doi.org/10.3390/info14070372>

Academic Editor: Luis Martínez López

Received: 15 May 2023

Revised: 26 June 2023

Accepted: 27 June 2023

Published: 29 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

pruning of the combinatorial complexity of such relationships. Since this sort of information often integrates with decision-making systems, effective user interfaces must be developed with visualizations or other representations for presenting the knowledge discovered.

For this paper, association rule development is our main focus [14]. Associations correspond to correlations among data items which are represented in the form of rules composed of attribute–value conditions that occur with some significant correlations in a set of data [15]. Association rules are of the form antecedents A , and consequents C , $A \rightarrow C$, where data entries satisfying the antecedent A commonly occur in conjunction with data values for C . The term “common occurrence” has a probabilistic implication and is not a functional dependency as used in databases.

Finally, we provide an outline of the paper:

Section 2: Fuzzy set representations to be used are defined. A review of previous research on fuzzy data mining, especially research on association rule data mining using fuzzy sets, is presented.

Section 3: The metrics used in association data mining are defined. Specifics of basic association rule data mining are given and a running example illustrates this. A case analysis of the bounds of the basic measures, support and confidence, and the interestingness metrics, lift and conviction, are provided.

Section 4: Next, the basics of the Apriori algorithm are illustrated using the running example.

Section 5: The extensions of association mining needed for intuitionistic fuzzy sets are described, in particular the required set cardinality used for association rule generation. The running example extended to fuzzy intuitionistic data is then used to illustrate the approach using the various metrics.

2. Background

In this section we provide an overview of the fuzzy set representations that are relevant to this paper. Then we discuss data mining in general and specifically fuzzy data mining. In particular, we review other research for fuzzy association rules.

2.1. Uncertainty Representations

In this section we briefly overview common uncertainty representations [16], including fuzzy sets and intuitionistic fuzzy sets for approaches to data mining.

2.1.1. Fuzzy Set Theory

Fuzzy set representations [17,18] provide the membership degrees of data values in a set, as opposed to crisp sets. For domain D , a fuzzy set, FS , is

$$FS(D) = \{ \langle a_i, m(a_i) \rangle \mid 0 \leq m(a_i) \leq 1, a_i \in D, i = 1 \dots n \}$$

where a_i is a data value and $m(a_i)$ is the membership of the data value.

2.1.2. Intuitionistic Fuzzy Sets

Intuitionistic fuzzy set theory extends ordinary fuzzy set theory by allowing both positive and negative memberships to be specified. Recall that an ordinary fuzzy set $FS(D) = \{ \langle a_i, m(a_i) \rangle \}$ has only one membership value for a data element a_i . An intuitionistic fuzzy set $IFS(D)$, [19] allows both positive, $m_S(a_i)$, and negative membership values, $m_S^*(a_i)$.

$$IFS(D) = \{ \langle a_i, m_S(a_i), m_S^*(a_i) \rangle \mid a_i \in D \} \text{ where } m_S(a_i), m_S^*(a_i) \in [0, 1].$$

Specifically, the sum of the membership, $m_S(a_i)$, and non-membership, $m_S^*(a_i)$, is not necessarily one, then: $0 \leq m_S(a_i) + m_S^*(a_i) \leq 1$. Additionally the hesitation $h_S(a_i)$

$$h_S(a_i) = 1 - (m_S(a_i) + m_S^*(a_i))$$

is the degree of indeterminacy (hesitation).

2.1.3. Interval-Valued Fuzzy Sets

Interval values are used in many areas to capture the imprecision and uncertainty of data. We first provide the formalisms for interval arithmetic [20,21] as needed for interval valued fuzzy sets. We let D be the domain and intervals will be represented by the values of the lower bound, $z_{\downarrow} = lb(a_i)$, and an upper bound, $z^{\uparrow} = ub(a_i)$, of an interval $I(a_i)$, for the data value $a_i \in D$

$$I(a_i) = [z_{\downarrow}, z^{\uparrow}] = \{z \in D \mid z_{\downarrow} \leq z \leq z^{\uparrow}\}$$

Now, in an interval-based fuzzy set, $IVF(D)$, representation is based on using upper, $mu(a_i)$, and lower bounds, $ml(a_i)$, on fuzzy memberships

$$IVF(D) = \{ \langle a_i, I(a_i) \rangle \mid I(a_i) = [ml(a_i), mu(a_i)] \}$$

For an interval $I(a_i)$, the size or length of the interval, IW , is just the difference of the lower and upper bounds,

$$IW(I(a_i)) = |ml(a_i) - mu(a_i)|.$$

IW is often used as a representation of the uncertainty of a data value a_i in an IVF as an information measure [22].

We note that IVF and IFS sets are equivalent as generalizations of fuzzy sets [23]. In particular, we have

$$m_S(a_i) = ml(a_i) \text{ and } m_S^*(a_i) = 1 - mu(a_i)$$

So,

$$IVF(D) = \{ \langle a_i, I(a_i) \rangle \mid I(a_i) = [m_S(a_i), 1 - m_S^*(a_i)] \}$$

This can be used for the set cardinality of $IVF(D)$, as related to the set cardinality of $IFS(D)$ developed for the calculation of support and confidence of fuzzy association rules.

2.2. Data Mining Approaches

We are primarily concerned with some of the particular algorithms used for knowledge discovery, but we will overview the complete processes involved in data mining. To begin, the initial basic steps involve data preparation [24]. First is a data cleaning stage, which includes resolution of missing data [25] or data errors. Additionally, the integration of data from multiple sources which may be heterogeneous is performed [26–28]. Next are the steps needed to prepare for actual data mining, which include the selection of the specific data relevant to the task and the transformation of this data into a format required by the data mining approach. Often, these steps are seen to be those involved in developing data warehouses. This provides an organization for the formatting of data to facilitate data mining approaches.

Knowledge discovery or data mining algorithms can generally be classified into two categories: predictive and descriptive data mining. The descriptive category includes association rules, classification, and class characterization. Typically, data generalizations or characterizations are provided for class descriptions, such as data summarizations [29]. Additionally, data class comparisons allow discrimination of classes to be developed.

Lastly, a classification approach uses data with known class memberships and then builds class models from the features extracted from the training data. Often, this is output as classification rules or a type of decision tree to support predictive classification of new data. Another well-developed knowledge discovery technique is predictive analysis with clustering approaches. Discovery of collections of data items that are similar is achieved by cluster analysis. A distance function is often developed by experts to provide effective similarity metrics. The characteristics of an appropriate clustering algorithm are that it results in high similarity for intra-cluster measures and low similarity for inter-clusters. To determine potential values that may be missing or the distribution of attribute

values, various prediction techniques may be used. Some such techniques include machine learning, genetic algorithms, and regression and correlation analysis.

2.3. Fuzzy Data Mining

Many early efforts in knowledge discovery appear in pattern recognition, especially in the form of fuzzy clustering [30]. There have been significant considerations of data mining involving uncertain and fuzzy data mining using neural networks and genetic algorithm approaches [2,3]. More specifically for our interest here in association rules, a number of approaches have been considered for fuzzy data [7,31,32].

2.3.1. Fuzzy Association Rules

In one approach, fuzzy association rule data mining has been developed which can use both fuzzy transactional and relational data [5]. Relational data enable extraction of multidimensional association rules and transactional data make the discovery of patterns more reliable. In another approach based on crisp sets of fuzzy transactions, a general model was developed to discover association rules [33]. The model was capable of being specialized to specific patterns or application data. Additionally, associative rules can be formed by using generalized implicit knowledge based on quantitative transaction data [34]. This technique uses a taxonomy, and items in the rules can be based on any level of the taxonomy. Another paper [35] proposed an automated method for autonomous mining of fuzzy association rules. They first find fuzzy sets by using an efficient clustering algorithm, and then determine their membership functions. Using these they find interesting fuzzy association rules. Two papers have proposed using type-2 fuzzy sets. In one, the quantitative values in transactions are dealt with as type-2 values. These can then be reduced to ordinary fuzzy data using split points and association data mining is carried out on this data [36]. Another approach [37] proposes a fuzzy frequent pattern-mining algorithm based on the type-2 fuzzy set theory of the data stream. The stream is partitioned based on a sliding window. This is then used to extract fuzzy association rules.

2.3.2. Fuzzy Spatial Association Rules

Some interesting approaches have focused on spatial data. Spatial association rules can be implemented by permitting spatial application features represented by fuzzy spatial objects and topological relationships [6]. They extract association rules using application-related spatial objects of interest and the fuzzy spatial features. Another approach to extracting association rules on spatially related data includes uncertain geographic and geologic information [38]. This makes use of fuzzy set cardinality to compute support and confidence metrics for rule evaluation.

None of these approaches have considered alternative uncertainty representations. In this paper we have extended the association rule approaches to consider the more flexible representation of uncertainty using intuitionistic fuzzy sets.

3. Association Rules

This section will describe the basis for the computation of association rules and provide examples demonstrating this. The extensions to the use of intuitionistic fuzzy sets will then be developed and the examples used to provide illustrations of the approach. A common example motivating association rules is the “market basket” of grocery items often purchased in the same transaction [14]. This sort of relationship can provide guidance on the marketing and placement of such items.

3.1. Association Rules Metrics

There are two main metrics, support and confidence, used for association rule data mining. Others are called interestingness metrics, used to evaluate discovered rules.

Let $E = \{e_1, \dots, e_n\}$ be a set of data items of interest. We will later consider the possibility that some of these items can involve uncertainty and be represented by intuitionistic

fuzzy values. For the data in E we examine a set of items, $\mathbf{R}$, that result from interactions such as a transaction consisting of items related to a purchase order or a query finding vacation locations for which desirable features appear. The interactions are specified by the resulting subsets of E , $\mathbf{R} = \{R_1, R_2, \dots, R_i, \dots\}$ where $R_i = \{e_j, \dots, e_k\} \subseteq E$.

Specifically, we are concerned with the possible relationships among the data items that occur in $\mathbf{R}$. Let $S_j, S_k \subseteq E$, then we are interested in rules of the form F_{jk}

$$F_{jk}: S_j \Rightarrow S_k$$

where S_j is the antecedent and S_k the consequent term of the rule. Such a rule means items in S_j co-occur with items in S_k . We must make an assessment of the value of such rules by using two common metrics, support and confidence, to be evaluated to quantify the relationships. We first define two counts, N_{sp} and N_{ant} , used in the computation of the support and confidence metrics.

$$N_{sp} = |\{R_i \in \mathbf{R} \mid S_j \cup S_k \subseteq R_i\}|$$

$$N_{ant} = |\{R_i \in \mathbf{R} \mid S_j \subseteq R_i\}|$$

So N_{sp} is a count of the number of rules of the form F_{jk} and N_{ant} is the number of occurrences of the rule antecedent S_j .

3.1.1. Support Metric: M_{sp}

The support of a rule is a measure of how frequently the antecedent, S_j , and consequent, S_k , of a rule appear in the same entry R_i of the result set $\mathbf{R}$. Using the count N_{sp} , we have

$$\text{Support: } M_{sp} = N_{sp} / |\mathbf{R}|$$

3.1.2. Confidence Metric: M_{cf}

The confidence factor indicates how strongly a relationship is represented in the set of results $\mathbf{R}$. That is if $S_j \subseteq R_i$ then it is also the case that $S_k \subseteq R_i$. So we use the count N_{ant} here.

$$\text{Confidence: } M_{cf} = N_{sp} / N_{ant}$$

The percentage of results R_i in $\mathbf{R}$ satisfying M_{cf} will determine if the rule is of value for decision-making.

The support and confidence can be interpreted as probabilities where we have

$$M_{sp}: \text{Prob}(S_j \cup S_k); M_{cf}: \text{Prob}(S_k | S_j)$$

Support and confidence metrics are distinct. Confidence measures a rule's strength while support is a measurement of statistical significance. In other words, we want a rule with enough support above a threshold, or else it is not representative enough of the data. Rules with support above a threshold are termed as being frequent.

3.2. Examples of Rule Support and Confidence

We can examine a simple example of responses in a query to illustrate the formation of association rules F_{jk} . So we consider a website that can be queried to locate desirable eco-tourism vacation locations. In Table 1 are responses to queries on vacation preferences related to various available activities which constitute the items of interest, such as camping, fishing, etc.:

$$E = \{S_1: \text{camp}; S_2: \text{fish}; S_3: \text{hike}; S_4: \text{raft}; S_5: \text{ski}\}.$$

Table 1. Vacation Features Items.

Query Responses	Location Features
R ₁	{S ₁ , S ₂ , S ₅ }: camp, fish, ski
R ₂	{S ₁ , S ₂ , S ₄ }: camp, fish, raft
R ₃	{S ₁ , S ₂ , S ₄ }: camp, fish, raft
R ₄	{S ₁ , S ₃ }: camp, hike
R ₅	{S ₁ , S ₃ , S ₄ }: camp, hike, raft
R ₆	{S ₂ , S ₃ , S ₅ }: fish, hike, ski
R ₇	{S ₁ , S ₂ , S ₃ }: camp, fish, hike

Analyzing these query results, we can obtain association rules and the corresponding support and confidence values. Table 2 shows an example of some representative potential rules relating activities to be evaluated. Note that, in general, all possible combinations of the items in E must be considered, which is then a potentially large space. We discuss the Apriori algorithm in the next section to help prune computation of these combinations.

Table 2. Rules for Results of Table 1.

Rules: F _{ijk}	M _{sp} —Support	M _{cf} —Confidence
1. F ₁₂ : camp → fish	4/7 = 0.57	4/6 = 0.66
2. F ₂₁ : fish → camp	4/7 = 0.57	4/5 = 0.8
3. F ₃₁ : hike → camp	3/7 = 0.43	3/4 = 0.75
4. F ₂₅ : fish → ski	2/7 = 0.28	2/5 = 0.4
5. F ₅₂ : ski → fish	2/7 = 0.28	2/2 = 1
6. F ₅₄ : ski → raft	0/7 = 0	0/2 = 0
7. F _{{12}4} : camp, fish → raft	2/7 = 0.28	2/4 = 0.5
8. F _{5{12}} : ski → camp, fish	1/7 = 0.14	1/2 = 0.5

Consider the first line in Table 2 for the rule F₁₂: “camp → fish”, where camp is the rule antecedent and fish the consequent. This rule shows that available locations with camping often occurred in conjunction with a location in which fishing was allowed. Then, for this rule, the support is 57% since camp and fish co-occur in four entries in Table 1, (R₁, R₂, R₃, R₆), out of the seven query responses. Furthermore, the confidence is 0.66% since the feature camp occurs in a total of six entries of Table 1. Additionally, the second rule, F₂₁, has 80% confidence since the feature S₂—fish appears in five of the responses. Note that not all potential rules have support such as F₅₄, since the items ski and raft do not co-occur in any of the responses.

In considering the rules in Table 2, we can see a range of support values. Most often we are interested in rules with support above a minimum threshold Th_{sp}. If the support were not high enough, then the rule would not be significant enough to be considered, or even simply not as preferred. So if a threshold were Th_{sp} = 40%, then the first three rules would be considered to be of interest. These three also have good confidence values, and so are strong rules worth consideration. We see that rules that have high confidence are not necessarily of interest as their support may be low (for example F₂₅ or F_{5{12}}).

3.3. Interestingness Metrics

Support and confidence are the most common metrics for association rules; however we will examine others that can provide alternative analyses for rules. The commonly used lift measure relates the rule confidence and the expected confidence [39]. Strong rules with high support and confidence can be uninteresting depending on how much

the rule antecedent and consequent are related. Lift is used to determine the statistical relationship between the antecedent and consequent, which relates to the interestingness or usefulness of the rule. Another metric is conviction, which relies on the expected support that the antecedent of the rule appears without the consequent. For these metrics, we obtain evaluations as to the prediction significance of the rule.

Although interestingness metrics are useful in evaluation of rules after they are determined, the rules were dependent on the thresholds for confidence and support. Similarly to decisions on fuzzy set memberships, this can be somewhat arbitrary and user-dependent, making consistency of the rule discovery process a problem. One interesting approach addressing this that has been proposed [40] can be applied if rule antecedents follow a determined distribution. Then the minimum confidence and support thresholds can be set based on using a one-sided statistical confidence interval with hypothesis testing.

3.3.1. Lift Metric

First we consider lift given as:

$$\text{Lift} = \text{confidence} / \text{expected confidence}$$

We must first specify the support of the consequent of F_{jk} by defining

$$N_{\text{con}} = |\{R_i \in \mathbf{R} \mid S_k \subseteq R_i\}|;$$

and so the support of the consequent is

$$\text{Consequent Support: } M_{\text{con}} = N_{\text{con}} / |\mathbf{R}|$$

Now the expected confidence for a rule is the product of the rule support and consequent support divided by the rule support:

$$(M_{\text{sp}} * M_{\text{con}}) / M_{\text{sp}} = M_{\text{con}}.$$

Then finally we obtain:

$$\text{Lift} = M_{\text{cf}} / M_{\text{con}} = (N_{\text{sp}} / N_{\text{ant}}) * (|\mathbf{R}| / N_{\text{con}})$$

For the rule $S_j \Rightarrow S_k$, lift is interpreted as the correlation of the S_j and S_k .

- (a) Lift > 1, positive correlation.
- (b) Lift < 1, negative correlation.
- (c) Lift = 1, correlation is independent.

So lift determines how much more frequently S_j and S_k appear concurrently than would be expected if they were independent statistically.

In our example rules in Table 3, four rules have a lift less than 1. For example in rules F_{12} and F_{21} there is not a significant correlation between the items S_2 : fish and S_1 : camp, as camp occurs without fish twice and fish occurs without camp once. Now consider rules F_{25} and F_{52} for which the lift is 1.4. There is a positive correlation as there is a strong prediction of S_2 : fish when S_5 : ski occurs since in the only two occurrences of ski, fish also appears in the response.

3.3.2. Conviction Metric

Next we examine the conviction metric

$$\text{Conviction} = (1 - M_{\text{con}}) / (1 - M_{\text{cf}})$$

Conviction is sensitive to the direction of rules, $S_j \Rightarrow S_k$ versus $S_k \Rightarrow S_j$. So, we see rule 1 (F_{12}) and rule 2 (F_{21}) differ in their direction, and conviction for rule 1 is 1.26 but for rule 2 it is 2.15. This can be attributed to the lower confidence (0.66) of rule 1, and so rule 2 has

greater interest by this measure, leading to the stronger conviction. For logical implications where confidence is 1, the value is unbounded ($+\infty$), as is true for rule 5 (F_{52}). That is, S_5 (ski) uniquely appears with S_2 (fish) and not vice-versa.

Table 3. Lift Values.

Lift < 1	Lift > 1
F_{12} : 0.93	F_{25} : 1.4
F_{21} : 0.93	F_{52} : 1.4
F_{31} : 0.87	$F_{\{12\}4}$: 1.12
$F_{5\{12\}}$: 0.88	

3.4. Case Analysis of Metrics

We can examine the range of values of the metrics discussed by an analysis of the possible extreme values. Let $|\mathbf{R}| = Z$ and consider a supported rule of interest F_{jk} . Then for F_{jk} , $\{S_j \cup S_k\}$ must appear at least once, but can in the extreme, however unlikely, appear as many as the maximum of Z times or occurrences. The range of counts for such rules is then

$$1 \leq N_{sp} \leq Z$$

3.4.1. Support and Confidence Analysis

So, the range for support, $M_{sp} = N_{sp}/Z$, is

$$1/Z \leq M_{sp} \leq Z/Z = 1$$

Since, for each occurrence of a rule F_{jk} the antecedent S_j must appear, its ranges are similarly

$$1 \leq N_{ant} \leq Z$$

Recall that confidence is $M_{cf} = N_{sp}/N_{ant}$. We analyze the possible cases in Table 4.

Table 4. Confidence Ranges.

	$N_{sp} = 1$	$N_{sp} = Z$
$N_{ant} = 1$	$M_{cf} = 1$	$\emptyset$ (not possible)
$N_{ant} = Z$	$M_{cf} = 1/Z$	$M_{cf} = Z/Z = 1$

Hence, the range of M_{cf} is:

$$1/Z \leq M_{cf} \leq 1/1 = 1$$

3.4.2. Lift Analysis

Now, we can perform a similar extreme case analysis for lift. Consider the relationship of N_{ant} and N_{con} . For $N_{sp} = 1$, the antecedent, S_j , and consequent, S_k , must both appear at least once, and independently can each appear up to the extreme of Z times. For $N_{sp} = Z$, they must both appear Z times. The analysis of the cases for the range of lift is in Table 5.

From Table 5, the range for lift is:

$$1/Z \leq \text{Lift} \leq Z$$

Table 5. Range of values for Lift: $(N_{sp}/N_{ant}) * (|R|/N_{con})$.

N_{ant}	N_{con}	Lift	Lift
		$N_{sp} = 1$	$N_{sp} = z$
1	1	$1 * Z / 1 * 1 = Z$	$\emptyset$ (not possible)
1	Z	$1 * Z / 1 * Z = 1$	$\emptyset$ (not possible)
Z	1	$1 * Z / Z * 1 = 1$	$\emptyset$ (not possible)
Z	Z	$1 * Z / Z * Z = 1/Z$	$Z * Z / Z * Z = 1$

3.4.3. Conviction Analysis

Finally, we examine conviction

$$\text{Conviction} = (1 - M_{con}) / (1 - M_{cf})$$

As for the antecedent, $1 \leq N_{con} \leq Z$. Then, since $M_{con} = N_{con}/Z$

$$1/Z \leq M_{con} \leq 1$$

Since $1/Z \leq M_{cf} \leq 1$ and $1/Z \leq M_{con} \leq 1$, the ranges for both the numerator and denominator are: $[1 - 1/Z, 1 - 1] = [(Z - 1)/Z, 0]$. However M_{con} and M_{cf} cannot both be 1, so there is not a 0/0 case. Then, for conviction the range is:

$$0 \leq \text{Conviction} \leq +\infty$$

4. Apriori Procedure

As discussed, a major issue for finding association rules is the combinatorial complexity of computing all potential appropriate combinations of data items. This issue has been addressed by developing algorithms using the Apriori property [41]. Examples of the approach are described in this section.

An overview of the overall approach for discovering strong association rules can be formulated in three stages:

- Computing the frequent item sets: This is performed using the support metric for evaluation and utilizing the Apriori property to simplify the search.
- Determining strong association rules: From the frequent item sets in the first stage, the confidence metric is used in the evaluation to determine strong rules.
- Evaluating effectiveness of the resulting strong rules: Interestingness metrics, such as lift and conviction, are used in the selection of the most useful strong rules.

We will illustrate some possible frequent item sets using an example from the recreation locations application. A major issue is that there is a combinatorially large number of possible frequent item sets. The Apriori algorithm is commonly used to reduce computational complexity. It makes use of the prior knowledge of the support frequency to help reduce the generation of frequent sets. This property means that all subsets of a frequent set are frequent and if a set is not frequent all of its supersets are not frequent. The basic characteristic of the Apriori property is that it treats the support measure as being anti-monotonic.

Apriori Example of Frequent Set Generation

Finding frequent item sets is an iterative process that determines the next candidate frequent set from previous ones. The search generates the next candidate $k + 1$ level item set from the co-join of the previous k level item set. The Apriori property is then used to prune the search by eliminating possible sets from the generated set. The process follows as all item supersets of any item set that is not frequent are not frequent and can be eliminated. This terminates when no new frequent sets are generated.

We illustrate this with data similar to the previous discussion, with the same item set,

$$E = \{S_1: \text{camp}; S_2: \text{fish}; S_3: \text{hike}; S_4: \text{raft}; S_5: \text{ski}\},$$

but suitable to illustrate the approach for computing frequent sets, Table 6.

Table 6. Example results for Apriori process.

Query Responses	Location Features
R ₁	{S ₁ , S ₂ , S ₄ }: camp, fish, raft
R ₂	{S ₂ , S ₅ }: fish, ski
R ₃	{S ₂ , S ₃ }: fish, hike
R ₄	{S ₁ , S ₃ }: camp, hike
R ₅	{S ₁ , S ₄ , S ₅ }: camp, fish, raft, ski
R ₆	{S ₂ , S ₃ }: fish, hike,
R ₇	{S ₁ , S ₃ }: camp, hike
R ₈	{S ₁ , S ₂ , S ₃ , S ₄ }: camp, fish, hike, raft
R ₉	{S ₁ , S ₂ , S ₃ }: camp, fish, hike

To start the algorithm, we determine the support for each of the five features, as their support values determine the frequent feature sets.

For simplicity, we will use the first letter of the features for the rest of this example.

The iterative process starts using Table 7 as the candidate set C₁ and generates the next possible set C₂ by a join, (C₁ $\oplus$ C₁), of the items for the next set of candidates:

$$\{CF, CH, CS, CR, FH, FS, FR, HS, HR, SR\}$$

Table 7. Example features.

Feature	Support
S ₁ : Camp	6/9–0.66
S ₂ : Fish	7/9–0.77
S ₃ : Hike	6/9–0.66
S ₄ : Raft	3/9–0.33
S ₅ : Ski	2/9–0.22

Then we apply the Apriori condition using Table 7, where a set is pruned if any of its subsets are not above the support count threshold. For example, the set CH has C and H as subsets both with 0.66 support above the threshold and so CH is not pruned. However, if our support threshold Th_{sp} is 0.3 then any set with S as a subset will be pruned.

Again, any sets with S as a subset, {CS, HS, FS, RS}, are pruned, since S is not frequent, and we obtain as the next candidate set, C₂, Table 8.

From this candidate set we generate the next stage by joining C₂ and C₂ giving:

$$\{CFH, CFR, CHR, FHR\}$$

Now, these must be checked to see if all subsets are frequent. For example, the subsets of {CFH} are {CF, CH, FH}, all of which are frequent. For {FHR} the subsets are {FH, FR, HR}, and HR is not frequent, and it is pruned. Similarly, CHR is pruned. So, C₃ = {CFH, CFR} and the join of these is {CFHR}. Since FHR is a non-frequent subset we go no further and the final result is C₃.

Table 8. Candidate set C_2 .

Item-Set	Support
$S_1 \oplus S_2$: CF	4/9–0.44
$S_1 \oplus S_3$: CH	4/9–0.44
$S_1 \oplus S_4$: CR	3/9–0.33
$S_2 \oplus S_3$: FH	4/9–0.44
$S_2 \oplus S_4$: FR	3/9–0.33
$S_3 \oplus S_4$: HR	1/9–0.11

5. Uncertainty Querying

In this section extensions to the use of intuitionistic fuzzy sets for association rules are developed based on extensions of the metrics. The examples used to provide illustrations of the approach are based on the running example of the querying of vacation locations.

We first examine more closely our example of querying of vacation locations described in Table 1. As well as the specific features for each site, there are usually criteria such as distance, costs, and location desirability influencing the user evaluation. With these subjective criteria specified by linguistic values such as lower cost, nearness, or pleasantness of location, we would then obtain a degree of the suitability of each location retrieved on this basis. We can use the term “desirable” for the combination of the factors and capture this using a fuzzy set representation for membership in the set of desirable locations. Basic fuzzy sets were used in previous research [38], and here we want to extend this to more flexible representations, including intuitionistic-based fuzzy memberships.

5.1. Fuzzy Intuitionistic Measures for Support and Confidence

Next we must consider how to deal with the uncertainty using intuitionistic fuzzy set representations of subjective results of an operation such as a query. In the cases here for a query response, R_i , we consider an intuitionistic fuzzy membership in $\mathbf{R}$ for the query.

$$\{ \langle R_i = \{e_j, \dots, e_k\}, m_{\mathbf{R}}(R_i), m^*_{\mathbf{R}}(R_i) \rangle \}$$

To determine frequent feature sets it is necessary to calculate the number of responses R_i that support the rule F_{jk} . For crisp operations, the support count N_{sp} is simply the size or cardinality of the set R_i . However, since we have a fuzzy membership, we must adapt the support count. For the case of a query with simple fuzzy membership, this was modeled using the set cardinality for fuzzy sets [42] as the sigma count, which is simply the sum of its membership values.

$$\text{Card}(\text{FS}(D)) = \sum_{i=1}^n m(a_i), a_i \in D$$

5.1.1. Cardinality of Intuitionistic Fuzzy Sets

The cardinality of intuitionistic-valued fuzzy sets (IFS) is needed in the procedure of determining association rules [43,44]. The development of the cardinality of an IFS set follows a geometrical framework [45]. Specifically, they define two bounds, the least cardinality:

$$\text{Min Card}(\text{IFS}(D)) = \sum_{i=1}^n m_S(a_i)$$

and the maximum cardinality:

$$\begin{aligned} \text{Max Card}(\text{IFS}(D)) &= \sum_{i=1}^n (m_S(a_i) + h_S(a_i)) = \sum_{i=1}^n (m_S(a_i) + 1 - (m_S(a_i) + m^*_S(a_i))) \\ &= \sum_{i=1}^n (1 - m^*_S(a_i)) \end{aligned}$$

Then, the overall cardinality is the interval of the least and maximum values:

$$\text{Card}(\text{IFS}(D)) = [\text{Min Card}(\text{IFS}(D)), \text{Max Card}(\text{IFS}(D))]$$

An interval may not be as convenient to use in calculations, so the average possibility cardinality AVCard can also be used:

$$\text{AVcard}(\text{IFS}(D)) = 1/2 (\sum_{i=1}^n (m_S(a_i) + 1 - m_{*S}(a_i))) = 1/2 (\sum_{i=1}^n (m_S(a_i) + m_S(a_i) + h_S(a_i))) \\ = \sum_{i=1}^n (m_S(a_i) + h_S(a_i))/2$$

It is noted that this is the midpoint of the interval of the Min–Max cardinalities for the IFS.

5.1.2. Intuitionistic Metrics

Now, we can develop the support and confidence measures for fuzzy intuitionistic memberships using these cardinality approaches. First adapting the counts, FN_{sp} , FN_{ant} , FN_{con} , for a rule F_{jk} we must define for each count, an index set.

$$I(\text{sp}) = \{i \mid S_j \cup S_k \subseteq R_i\}; I(\text{ant}) = \{i \mid S_j \subseteq R_i\}; I(\text{con}) = \{i \mid S_k \subseteq R_i\};$$

Since the cardinality we use here is an interval, there are two measures for the Min and Max cardinality, respectively, in the counts:

$$\text{FN}_{\text{sp}} = |\{ \langle R_i, m_R(R_i), m_{*R}(R_i) \rangle \mid S_j \cup S_k \subseteq R_i \}|$$

$$\text{MinFN}_{\text{sp}} = \text{Min}|\{ \langle R_i, m_R(R_i), m_{*R}(R_i) \rangle \}| = \sum_{i \in I(\text{sp})} m_R(R_i)$$

$$\text{MaxFN}_{\text{sp}} = \text{Max}|\{ \langle R_i, m_R(R_i), m_{*R}(R_i) \rangle \}| = \sum_{i \in I(\text{sp})} (m_R(R_i) + h_R(R_i)) = \sum_{i \in I(\text{sp})} (1 - m_{*R}(R_i))$$

Similarly, for antecedent and consequent counts:

$$\text{MinFN}_{\text{ant}} = \sum_{i \in I(\text{ant})} m_R(R_i); \text{MaxFN}_{\text{ant}} = \sum_{i \in I(\text{ant})} (1 - m_{*R}(R_i)) \\ \text{MinFN}_{\text{con}} = \sum_{i \in I(\text{con})} m_R(R_i); \text{MaxFN}_{\text{con}} = \sum_{i \in I(\text{con})} (1 - m_{*R}(R_i))$$

Now the measures become two values for support and confidence each. First for fuzzy supports:

$$\text{MinFM}_{\text{sp}} = \text{MinFN}_{\text{sp}} / |\mathbf{R}|$$

$$\text{MaxFM}_{\text{sp}} = \text{MaxFN}_{\text{sp}} / |\mathbf{R}|$$

Following a similar approach for the counts for the antecedent, we have a min and a max count. Then, to determine fuzzy confidence, we use::

$$\text{MinFM}_{\text{cf}} = \text{MinFN}_{\text{p}} / \text{MinFN}_{\text{ant}}$$

$$\text{MaxFM}_{\text{cf}} = \text{MaxFN}_{\text{sp}} / \text{MaxFN}_{\text{ant}}$$

However, to proceed, we must consider that there are two possible interpretations of $|\mathbf{R}|$. It is an issue of whether or not the size or cardinality of $\mathbf{R}$ should be based on the membership values. So we will use two interpretations of $|\mathbf{R}|$:

$$1. \text{Min}|\mathbf{R}| = \sum_{i=1}^n m_R(R_i) \quad 2. \text{Max}|\mathbf{R}| = \sum_{i=1}^n (1 - m_{*R}(R_i))$$

5.2. Fuzzy Query Example

We will first show the results of the query using the example of Table 1 with the example fuzzy memberships. Then we compute support and confidence using the extended measures we have developed.

Next, using the data in Table 9, we evaluate support and confidence for the same example rules as shown in Table 2.

Table 9. Query responses with intuitionistic uncertainty.

Query Responses	Location Features	Intuitionistic Membership (m, m*)
R ₁	{S ₁ , S ₂ , S ₅ }: camp, fish, ski	<0.6, 0.3>
R ₂	{S ₁ , S ₂ , S ₄ }: camp, fish, raft	<0.5, 0.3>
R ₃	{S ₁ , S ₂ , S ₄ }: camp, fish, raft	<0.8, 0.2>
R ₄	{S ₁ , S ₃ }: camp, hike	<0.6, 0.4>
R ₅	{S ₁ , S ₃ , S ₄ }: camp, hike, raft	<0.9, 0.1>
R ₆	{S ₂ , S ₃ , S ₅ }: fish, hike, ski	<0.8, 0.1>
R ₇	{S ₁ , S ₂ , S ₃ }: camp, fish, hike	<0.7, 0.2>

5.3. Discussion of Results

It is clear that the most consistent use of $|R|$ is to use the min value for the min support and the max value for the max support. These results for Min/Min and Max/Max produce values bracketing the support values for the crisp cases in Table 2. Averaging these produces values about 1% different from the crisp support values. Additionally, the confidence values are compatible with previous values. This shows that we can use the intuitionistic fuzzy sets to provide flexibility in capturing uncertainty in analyzing data and obtain useful results, Table 10.

Table 10. Support and Confidence.

Rules: F _{jk}	MinFM _{sp} Min R	MinFM _{sp} Max R	MaxFM _{sp} Min R	MaxFM _{sp} Max R	MinFM _{cf}	MaxFM _{cf}
1. F ₁₂ : camp → fish	0.529	0.479	0.615	0.557	0.63	0.67
2. F ₂₁ : fish → camp	0.529	0.479	0.600	0.557	0.76	0.77
3. F ₃₁ : hike → camp	0.449	0.402	0.472	0.428	0.76	0.74
4. F ₂₅ : fish → ski	0.286	0.259	0.329	0.298	0.41	0.41
5. F ₅₂ : ski → fish	0.286	0.259	0.329	0.298	1.0	1.0
6. F ₅₄ : ski → raft	0.0	0.0	0.0	0.0	0.0	0.0
7. F _{{12}4} : camp, fish → raft	0.271	0.246	0.30	0.272	1.0	1.0
8. F _{5{12}} : ski → camp, fish	0.129	0.117	0.143	0.129	0.43	0.44

5.3.1. Effect of Negative Memberships

In our example of intuitionistic data in Table 9 we chose to focus only on stronger positive memberships to allow a consistent analysis of rules. Now we examine the effect if a few data values have higher negative memberships. Specifically, we will use a membership of <0.3, 0.6> for R₆ and then R₁ to compare the effects on rule support. For the case of R₆, we consider rules F₂₅ and F₅₂ as there is only one other entry, R₁, involved. We expect less significant changes for rules in which more data is involved, and the effect of the one change would not be expected to be as strong.

For rules F₂₅ and F₅₂, the change affects $|R|$ and FN_{sp}:

$$\text{Min } |R| = 4.4; \text{Max } |R| = 4.9; \text{Min FN}_{\text{sp}} = 0.9; \text{Max FN}_{\text{sp}} = 1.1$$

So,

$$\text{Min FM}_{\text{sp}} = 0.9/4.4 = 0.205; \text{Max FM}_{\text{sp}} = 1.1/4.9 = 0.224$$

Then, the support for both rules has decreased, to a support of 28% for MinFM_{sp} and 25% for Max FM_{sp}.

Next, we replace the membership of R_1 with $\langle 0.3, 0.6 \rangle$ and compare the effect on F_{12} . Here,

$$\text{Min } |\mathbf{R}| = 4.6; \text{Max } |\mathbf{R}| = 5.1; \text{Min } \text{FN}_{\text{sp}} = 2.3; \text{Max } \text{FN}_{\text{sp}} = 2.7$$

Then,

$$\text{Min } \text{FM}_{\text{sp}} = 2.3/4.6 = 0.5; \text{Max } \text{FM}_{\text{sp}} = 2.7/5.1 = 0.529$$

As expected, the changes in the support are now much less—roughly 9.5% for both supports.

5.3.2. Lift Metric

Now, we examine the lift metric with the data of Table 9. Again, the results are consistent with the lift results of Table 3, as the same three rules, F_{25} , F_{52} , $F_{\{12\}4}$, have lift >1 , Table 11.

Finally, to be complete, we examine the same change in negative membership for R_6 and R_1 above. For F_{25} and F_{52} , we see:

$$\text{Lift Min: } 1.52\text{--}6\% \text{ increase; Lift Max: } 1.44\text{--}3.5\% \text{ increase.}$$

For F_{12} , the lift value changes are marginal:

$$\text{Lift Min: } 0.90\text{--}1\% \text{ decrease; Lift Max: } 0.91\text{--}1\% \text{ decrease.}$$

These results are still consistent with previous values, since F_{25} and F_{52} have lift >1 , and F_{12} has lift <1 . The increases for F_{25} and F_{52} reflect the change in the co-occurrence of the antecedents and consequents due to the increased negative membership as shown in Section 5.3.1 above. Additionally, for F_{12} the change is negligible, again consistent with the smaller changes in its support for similar reasons.

Table 11. Lift values for Intuitionistic Memberships.

Rules: F_{jk}	Lift Min	Lift Max
1. F_{12} : camp $\rightarrow$ fish	0.91	0.92
2. F_{21} : fish $\rightarrow$ camp	0.9	0.92
3. F_{31} : hike $\rightarrow$ camp	0.9	0.89
4. F_{25} : fish $\rightarrow$ ski	1.43	1.39
5. F_{52} : ski $\rightarrow$ fish	1.44	1.38
6. F_{54} : ski $\rightarrow$ raft	0	0
7. $F_{\{12\}4}$: camp, fish $\rightarrow$ raft	1.1	1.2
8. $F_{5\{12\}}$: ski $\rightarrow$ camp, fish	0.81	0.79

6. Conclusions

We have shown that it is quite feasible to discover association rules appropriate for applications in which the data involves intuitionistic fuzzy set descriptions.

The extensions required were applications of the set cardinality for intuitionistic fuzzy sets for determining the counts used in the metrics. This produced consistent results for the support and confidence of rules. We also examined the effects of data with higher negative memberships in some cases. The results implied that the approaches were still valid, depending on the percentage of data with such memberships. Additionally, to provide insight into the metrics used in the approach, the ranges of support and confidence as well as lift and conviction were determined.

Since interval valued fuzzy sets have similar formulations, the use of such uncertainty measures can be considered next for association rules. Entropy measures can be used [22] to evaluate the IFS and IVFS in the association rule computations.

Author Contributions: Conceptualization: F.P. and R.Y.; Methodology: F.P. and R.Y.; Writing: F.P. and R.Y. All authors have read and agreed to the published version of the manuscript.

Funding: Fred Petry would like to acknowledge the Naval Research Laboratory's Base Program for sponsoring this research.

Data Availability Statement: No new data were created for this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Aggarwal, C.; Li, Y.; Wang, J.; Wang, J. Frequent pattern mining with uncertain data. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 29–38. [\[CrossRef\]](#)
- Hirota, K.; Pedrycz, W. Fuzzy computing for data mining. *Proc. IEEE* **1999**, *87*, 1575–1599. [\[CrossRef\]](#)
- Mirzakhani, V. Value of fuzzy logic for data mining and machine learning: A case study. *Expert Syst. Appl.* **2020**, *162*, 113781. [\[CrossRef\]](#)
- Petry, F.; Yager, R. Intuitionistic and interval-valued fuzzy set representations for data mining. *Algorithms* **2022**, *15*, 249. [\[CrossRef\]](#)
- Au, W.; Chan, K. Mining fuzzy association rules in a bank-account database. *IEEE Trans. Fuzzy Syst.* **2003**, *11*, 238–248.
- da Silva, H.; Felix, T.; de Venâncio, P.; Carniel, A. Discovery of Spatial Association Rules from Fuzzy Spatial Data. In Proceedings of the Conceptual Modeling: 41st International Conference, ER 2022, Hyderabad, India, 17–20 October 2022; Ralyté, J., Chakravarthy, S., Mohania, M., Jeusfeld, M.A., Karlapalem, K., Eds.; Lecture Notes in Computer Science. Springer: Cham, Switzerland, 2022; Volume 13607.
- Lin, C.; Li, T.; Fournier Viger, P.; Hong, T. A fast algorithm for mining fuzzy frequent itemsets. *J. Intell. Fuzzy Syst.* **2015**, *29*, 2373–2379. [\[CrossRef\]](#)
- Szmidt, E.; Kacprzyk, J. Medical Diagnostic Reasoning Using a Similarity Measure for Intuitionistic Fuzzy Sets. In Proceedings of the Eighth International Conference on IFSS, Varna, Bulgaria, 20–21 June 2004; pp. 61–69.
- Dubois, D.; Gottwald, S.; Hajek, P.; Kacprzyk, J.; Prade, H. Terminological difficulties in fuzzy set theory—The case of intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **2005**, *156*, 485–492. [\[CrossRef\]](#)
- Solanki, S.; Patel, J. A survey on association rule mining. In Proceedings of the 2015 Fifth International Conference on Advanced Computing & Communication Technologies, Haryana, India, 21–22 February 2015; pp. 212–216. [\[CrossRef\]](#)
- Antonie, M.; Zaiiane, O. Mining positive and negative association rules: An approach for confined rules. In Proceedings of the Knowledge Discovery in Databases: PKDD 2004: 8th European Conference on Principles and Practice of Knowledge Discovery in Databases, Pisa, Italy, 20–24 September 2004; pp. 27–38.
- Dong, X.; Hao, F.; Zhao, L.; Zu, T. An efficient method for pruning redundant negative and positive association rules. *Neurocomputing* **2020**, *393*, 245–258. [\[CrossRef\]](#)
- Han, J.; Pei, J.; Tong, H. *Data Mining: Concepts and Techniques*, 4th ed.; Morgan Kaufmann: San Francisco, CA, USA, 2023.
- Ceglar, A.; Roddick, J. Association mining. *ACM Comput. Surv.* **2006**, *38*, 5:1–5:42. [\[CrossRef\]](#)
- Hipp, J.; Güntzer, U.; Nakhaeizadeh, G. Algorithms for association rule mining—A general survey and comparison. *SIGKDD Explor.* **2000**, *2*, 58–64. [\[CrossRef\]](#)
- Kruse, R.; Mostaghim, S.; Borgelt, C.; Braune, C.; Steinbrecher, M. *Computational Intelligence: A Methodological Introduction*, 3rd ed.; Springer Nature: Cham, Switzerland, 2022.
- Klir, G.; St. Clair, U.; Yuan, B. *Fuzzy Set Theory: Foundations and Applications*; Prentice Hall: Hoboken, NJ, USA, 1997.
- Zadeh, L. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [\[CrossRef\]](#)
- Atanassov, K. Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **1986**, *20*, 87–96. [\[CrossRef\]](#)
- Deschrijver, G. Arithmetic operators in interval-valued fuzzy set theory. *Inf. Sci.* **2007**, *177*, 2906–2924. [\[CrossRef\]](#)
- Moore, R.; Kearfott, B.; Cloud, M. *Introduction to Interval Analysis*; SIAM: Philadelphia, PA, USA, 2009.
- Burillo, P.; Bustince, H. Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. *Fuzzy Sets Syst.* **1996**, *78*, 305–316. [\[CrossRef\]](#)
- Atanassov, K.; Gargov, G. Interval valued intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **1986**, *31*, 343–349. [\[CrossRef\]](#)
- Rajagopalan, B.; Isken, M. Exploiting data preparation to enhance mining and knowledge discovery. *IEEE Trans. Syst. Man Cybernetics* **2001**, *31*, 460–467. [\[CrossRef\]](#)
- Islam, M.; Anderson, D.; Petry, F.; Smith, D.; Elmore, P. The fuzzy integral for missing data. In Proceedings of the 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, Italy, 9–12 July 2017.
- Elmore, P.; Anderson, D.; Petry, F. Evaluation of heterogeneous uncertain information fusion. *Ambient. Intell. Humaniz. Comput.* **2020**, *11*, 799–811. [\[CrossRef\]](#)
- Petry, F.; Elmore, P.; Yager, R. Combining uncertain information of differing modalities. *Inf. Sci.* **2015**, *322*, 237–256. [\[CrossRef\]](#)
- Yager, R.; Petry, F. An intelligent quality based approach to fusing multi-source probabilistic information. *Inf. Fusion* **2016**, *31*, 127–136. [\[CrossRef\]](#)
- Ahmed, M. Data summarization: A survey. *Knowl. Inf. Syst.* **2019**, *58*, 249–273. [\[CrossRef\]](#)
- Bezdek, J. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Plenum Press: New York, NY, USA, 1981.

31. Alam, M.; Ahmed, C.; Samiullah, M.; Leung, C. Mining frequent patterns from hypergraph databases. In Proceedings of the Advances in Knowledge Discovery and Data Mining: 25th Pacific-Asia Conference, PAKDD 2021, Virtual Event, 11–14 May 2021; Part II. pp. 3–15.
32. de Graaf, J.; Kusters, W.; Witteman, J. Interesting fuzzy association rules in quantitative databases. In *Principles of Data Mining and Knowledge Discovery*; LNAI 2168; Springer: Berlin, Germany, 2001; pp. 140–151.
33. Delgado, M.; Marin, N.; Sanchez, D.; Vila, M. Fuzzy association rules: General model and applications. *IEEE Trans. Fuzzy Syst.* **2003**, *11*, 214–225. [[CrossRef](#)]
34. Hong, T.; Lin, K.; Wang, S. Fuzzy data mining for interesting generalized association rules. *Fuzzy Sets Syst.* **2003**, *138*, 255–269. [[CrossRef](#)]
35. Kaya, M.; Alhajj, R.; Polat, F.; Arslan, A. Efficient automated mining of fuzzy association rules. In Proceedings of the Database and Expert Systems Applications: 13th International Conference, DEXA 2002, Aix-en-Provence, France, 2–6 September 2002; pp. 133–142.
36. Chen, C.; Hong, T.; Li, Y. Fuzzy association rule mining with type-2 membership functions. In Proceedings of the Asian Conference on Intelligent Information and Database Systems, Bali, Indonesia, 23–25 March 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 128–134.
37. Chen, J.; Lia, P.; Fanga, W.; Zhoua, N.; Yina, Y.; Xua, H.; Zheng, H. Fuzzy Association rules mining based on type-2 fuzzy sets over data stream. *Procedia Comput. Sci.* **2022**, *199*, 456–462. [[CrossRef](#)]
38. Ladner, R.; Petry, F.; Cobb, M. Fuzzy set approaches to spatial data mining of association rules. *Trans. GIS* **2003**, *7*, 123–138. [[CrossRef](#)]
39. Sael, N.; Alashqur, A.; Sowan, B. Using the interestingness measure lift to generate association rules. *J. Adv. Comput. Sci. Technol.* **2015**, *4*, 156–162.
40. Chen, S.; Tsai, T.; Chung, C.; Li, W. Dynamic association rules for gene expression data analysis. *BMC Genom.* **2015**, *16*, 786. [[CrossRef](#)]
41. Aggarwal, R.; Imielinski, T.; Swami, A. Mining Association Rules between sets of items in large databases. In Proceedings of the ACM-SIGMOD International Conference on Management of Data, Washington, DC, USA, 25–28 May 1993; ACM Press: New York, NY, USA, 1993; pp. 207–216.
42. Yen, J.; Langari, R. *Fuzzy Logic: Intelligence, Control and Information*; Prentice Hall: Upper Saddle River, NJ, USA, 1999.
43. Ioannis, K.; Vlachos, G.; Sergiadis, D. Subsethood, entropy, and cardinality for interval valued fuzzy sets—An algebraic derivation. *Fuzzy Sets Syst.* **2007**, *158*, 1384–1396.
44. Tripathya, B.; Jenab, S.; Ghosh, S. An intuitionistic fuzzy count and cardinality of intuitionistic fuzzy sets. *Malay J. Mat.* **2013**, *4*, 123–133. [[CrossRef](#)]
45. Szmidt, E.; Kacprzyk, J. Entropy for intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **2001**, *118*, 467–477. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.