Measuring Interestingness – Perspectives on Anomaly Detection

Adnan Masood[1*] Stephen Soong[2]

1. Graduate School of Computer and Information Sciences. 3301 College Avenue, Fort Lauderdale-Davie, Florida 33314-7796.
2. Cornell University, Ithaca, New York 14853

* E-mail of the corresponding author: adnan@nova.edu; ss24@cornell.edu

Abstract
We live in a data deluge. Our ability to gather, distribute, and store information has grown immensely over the past two decades. With this overabundance of data, the core knowledge discovery problem is no longer in the gathering of this data, but rather in the retrieving of relevant data efficiently. While the most common approach is to use rule interestingness to filter results of the association rule generation process, study of literature suggests that interestingness is difficult to define quantitatively and is best summarized as, *"a record or pattern is interesting if it suggests a change in an established model."* In this paper we elaborate on the term *interestingness*, and the surrounding taxonomy of interestingness measures, anomalies, novelty and surprisingness. We review and summarize the current state of literature surrounding interestingness and associated approaches.
Keywords: Interestingness, anomaly detection, rare-class mining, Interestingness measures, outliers, surprisingness, novelty

1. Introduction

From a machine learning perspective, the hypothesis behind term *interestingness* is arduous to formally define and quantify. Study of literature suggests that there is no agreement on formal definition of "interestingness"; this notion is best summarized as, *"record or pattern is interesting if it suggests a change in an established model."* This multi-disciplinary concept portrays interestingness as an entity that captures the impression of "novel" or "surprising". In search of the question "What's Interesting?", [1] attempts to answer by stating that *"Interestingness depends on the observer's current knowledge and computational abilities. Things are boring if either too much or too little is known about them, if they appear trivial or random."*

A similar multi-disciplinary construct like interestingness manifests the rare class entities in data and is often referred to as *anomaly*. Anomalies are data points or entities that do not agree with the expected model. In research literature, data mining and machine learning communities, the classification problem of outlier analysis and detection is often referred to with various different terminologies. As noted by Chandola [2] in their anomaly detection survey, it is cited as anomaly, novelty, chance discovery, exception mining, mining rare classes, and, informally, finding the needle in the haystack. Within the context of data mining, anomalies are the data points which not represented by the model, i.e. data points from a never before seen class. Similarly, in statistics, rare class entities are embodied as novelty, deviations, anomalies or outliers.

The machine learning areas of *interestingness* and *rare class mining* have a large body of academic work devoted although the distinction is often subjective; [3] notes *one man's outlier is another man's novelty*. Let's consider the definitions used by [4], Chandola [2], Markou [5] and Tan [6] respectively:

This process of retrieving the areas that are "interesting" for the understanding of the event is called "anomaly detection. [4],

"Novelty detection is the identification of new or unknown data or signals that a machine learning system is not aware of during training." [5]

"An outlier is an observation that differs so much from other observations as to arouse suspicion that it was generated by a different mechanism." [6]

2. An Overview of Interestingness Surveys

Few surveys of interestingness in production rules have been performed in the past by [7-9] and one most recently by [10]. In first survey in 1999 by [7] on "Knowledge Discovery and Interestingness Measures," the researchers examined an enumeration of 17 measures of rule interestingness, offering a brief description of each rule. These rules[1] range from Agrawal and Srikant's Item-set measures [11, 12], such as "interesting rules exceed a certain threshold of confidence $(P(B|A))$ and support $(P(AB))$ to more complex rules including Piatetsky-Shapiro's Rule-Interest Function [13]. Smyth and Goodman's J-Measure [14], Major and Mangano's Rule Refinement [15], Klemettinen et al. Rule Templates [16], Matheus and Piatetsky-Shapiro's Projected Savings [17], Hamilton and Fudger's I-Measures [18], Silbershatz and Tuzhilin's Interestingness [19], Kamber and Shinghal's Interestingness [20], Hamilton et al. Credibility Generalized [21], Liu et al. General Impressions [22], Gago and Bento's Distance Metric [23], Freitas' Surprisingness [24], Gray and Orlowska's Interestingness [25], Dong and Li's Interestingness [26], Liu et al. Reliable Exceptions [27] and Zhong et al. Peculiarity [28].

This assortment of objective and subjective measures of interest, commonly referred to as interestingness measures, is further classified as distance-based, probabilistic, or syntactic. [3] provides brief description of these measures as follows. Table 1 contains a detailed list of interestingness measures.

- **Piatetsky-Shapiro**: Deviation from statistical independence between the antecedent and the consequent: $(P(A|B) - P(A)P(B))$; the higher the deviation, the more interesting is the measure.
- **J-Measure: T**he average information content of a classification rule where given attributes are discrete valued,

$$P(AB)\log\frac{P(AB)}{P(B)} + P(A \sim B)\log\frac{P(\sim B|A)}{P(\sim B)}$$

The higher the J-values are, more interesting the measure is.

- **Gaga-Bento**: Distance metric to measure the distance between two rules, where distance is a function of the number of overlapping attributes common to two rules. Rules which are very distant from other rules are more interesting i.e. qualify to be outliers.
- **Zhong** - Peculiarity is a distance metric. In this case if the antecedents to a rule are similar to those of other rules, but its consequents are different, then the rule is interesting.
- **Silbershatz-Tuzhilin** - Measure of the extent to which a soft belief (hypothesis with "low" confidence) is changed in light of new evidence.
- **Freitas** - The explicit search for occurrences of Simpson's paradox, a seemingly self-contradictory statistical occurrence wherein conclusions drawn from a large data set are contradicted by conclusions drawn from subsets of the large data set.
- **Klemettin** - Rule templates are specified to identify the syntactic structure of either desired rules or undesired rules.

The survey performed by [7] provides a combination of both objective and subjective rules creating a good overall survey of researchers' efforts to define the interestingness of association rules. When Hilderman [29] reviewed the field again four years later, an additional 33 rules had been developed due to the field's growth.

The 2005 survey paper on Interestingness Measures for Knowledge Discovery [8] evaluated then-current research literature on the various techniques for determining the interestingness of patterns discovered by the data mining process. During the analysis, McGarry defines objective measures as those that are based upon the structure of the discovered patterns, while subjective measures are based upon user beliefs or biases regarding relationships in the data. This survey identifies the primary disadvantage of a subjective or

[1] Refer to Table 1 for detailed list of interestingness measures.

user-driven approach: that it limits the knowledge discovery process to user's hypothesis. In contrast, objective patterns are data-driven and therefore may manifest knowledge which is already known. This ultimately poses a research challenge to unify objective and subjective measures. The taxonomy of interestingness measures as noted by McGarry follows.

- Objective
  - Coverage
  - Support
  - Accuracy
- Subjective
  - Unexpected
  - Actionable
  - Novel

McGarry's [8] survey of interestingness measures for knowledge discovery approaches the topic in terms of data mining and knowledge discovery. Included in the paper as objective measures are standard statistical/information theoretic measures such as Shannon Entropy [30], Lorenz measure, Gini Index, Kullback-Leibler Distance, and the Atkinson Inequality, as well as the measures reviewed earlier by Hilderman [7, 29]. The term "distance" in this context is actually a measure of difference. None of the measures used are distance measures in the geometric sense. McGarry concludes with future research directions, primarily highlighting the strain between the objective and subjective approaches to finding interesting association rules. As discussed earlier in this paper regarding objective and subjective measure, McGarry states that subjective rules must necessarily constrain rule discovery to what a user expects to find and, consequently, unanticipated rules are indiscoverable. On the other hand, objective measures of interestingness will find rules that are of no interest to the user, since no context guides the discovery process. McGarry identified the resolution of this strain as an open question. A proposed solution is to find measures of interestingness, such as Simpson's Paradox detection explored by [31], that provide a middle ground to both approaches.

The subsequent notable and comprehensive survey was performed by [9] for interestingness measures in data mining. This survey identifies interestingness as a broader concept which constitutes of *conciseness, coverage, reliability, peculiarity, diversity, novelty, surprisingness, utility and actionability*. Bourassa [3] noted it as a very thorough review of interestingness measures and their properties. It distinguishes itself from McGarry's work in departing from a data mining context and instead focusing on measure categorization and behavior. The authors identify their research complimentary to McGarry's original work.

Geng and Hamilton [9] classified these interestingness measures based on the fundamental calculation or methodology for each measure (i.e., utilitarian, probabilistic, syntactic, distance). Majority of interestingness measures cited in Geng's survey are probabilistic in nature. Geng's review highlights the scope of the measures available to three types of rules: association, classification, and summaries (rule sets the paper reiterates the absence of a single definition for interestingness. Based on the diversity of measure definitions, the paper has compiled nine rule-interestingness criteria. They are as follows:

1. Conciseness: A pattern is concise if it contains few attribute-value pairs. A concise pattern is easy to understand, remember, and add to a user's knowledge (extends to sets of patterns).

2. Generality/Coverage: The generality or coverage of a pattern is a measure of how large a subset of the data the pattern covers. Patterns that characterize more information are interesting.

3. Reliability: a reliable pattern describes a relationship in the data that applies to a high percentage of the data.

4. Peculiarity: a pattern is peculiar if, by some distance measure, it lies far from other discovered

patterns.

5. Diversity: a pattern is diverse if it consists of elements that differ significantly from each other (extends to sets of patterns).

6. Novelty: a pattern is novel if has never been seen before and could not have been inferred from previously seen patterns;

7. Surprisingness (unexpectedness): the property of a pattern which contradicts existing knowledge or expectations.

8. Utility: the utility of a pattern is measured by its usefulness in reaching a goal (e.g. a business can use a sales pattern or market basket analysis to increase profits).

9. Actionability/Applicability: an actionable pattern enables decision making about future actions in a desired domain.

Geng then reviewed 38 objective, 3 subjective, and 2 semantic interestingness measures for association/classification rules according to the nine interestingness criteria. Since Geng's work [9], the most recent survey on knowledge discovery interestingness measures based on unexpectedness is by Kontonasios et al [10] which summarizes the primary features of syntactical and probabilistic approaches to interestingness mining.

3. Interestingness and Associated Taxonomy

The proposed definition of interestingness by [3] is as follows:

Data situated near the boundaries of models are interesting with a degree of interestingness inversely proportional to the distance from the boundaries of the models.

This definition, though in some ways an outgrowth from previous works by [7, 9, 10, 32, 33], differs from the current production-rule definition of interestingness in that it does not rely on descriptive measures of clusters, such as support or confidence. Instead, it proposes to approach rules that express clusters on the periphery of known models as interesting. Consequently, the proposed definition can capture the concepts of context and creativity for broader applicability beyond data mining. Further, this approach labels certain intuitively interesting points as interesting that other approaches would identify as outliers.

3.1 Hierarchy of Interestingness

This new definition entails the following hierarchy of interestingness, designed in order to align more closely with human intuition.

1. Explained: a point/cluster lying within the decision surface of a model describing a cluster.

2. Anomaly: a point/cluster lying near, but within, the decision surface of a model describing a cluster.

3. Interesting: a point/cluster lying on or beyond the decision surfaces of model describing a cluster, which is, lying between clusters.

4. Novel: a point/cluster lying beyond the decision surface of a model that encapsulates all known data (for example, the training data).

5. Noise: a point/cluster lying well beyond the decision surface of a model that encapsulates all known data.

This hierarchy provides three significant benefits. First, it guides the analyst to focus the most significant data for better-informed decision making; second, it provides data exploration direction by defining the boundaries of the model according to the location of interesting data records; third, it encourages meta-analysis of data records (through the application of metrics to the interestingness rankings of large amounts of data).

3.2 Criteria for Defining Interestingness

This approach also satisfies five criteria for the definition of interestingness better than the previous approaches do. Using Frawley's general definition of pattern where relevant [13], the definition aligns with the following criteria for the definition of interestingness:

1. The definition is not problem-specific. There is no assumed distribution, and the data defines the model. These characteristics allow the definition to be objective.

2. The definition is applicable to records or clusters. The definition is appropriate both to individual data points and clusters of points; unlike other approaches, it does not suppose all outliers to be of the same class.

3. The definition has geometric interpretation. It expresses the human experience in which the geometric location of a point conveys a sense of interestingness.

4. The definition identifies interesting records/ regions. The observation that a change in the observer's model signifies interestingness inspires the hierarchical classification (Explained, Anomaly, Interesting, Novel, Noise) outlined previously.

5. The definition captures human experience. The definition considers the work of Davis' and Berlyne in the study of the human experience of interestingness and accommodates notions of uncertainty, conflict, context, creativity, insight, etc.

3.3 Multi-disciplinary View of Interestingness

In constructing this definition, several perspectives on interestingness were considered, namely those of social science, creativity, association rules, information theory, and anomaly detection.

In social science, this approach heavily considers Davis' work, which comments on complications such as different audiences, changing audience assumptions, and the "timing" of a theory. From social science, the idea that interesting data is that which challenges the models that have been established to describe a data set is gained.

From creativity, clear parallels appear between epistemic novelty and the data mining process. This research considers the potential of online systems to "learn" from new data so that one can easily see where the frequentist notions of novelty can be employed. Particularly relevant are thee concepts of novelty outlined by Saunders that suggest, to some extent, means of identifying novel data; among them, uncertainty, conflict, surprise, and incongruity may all be labeled as novel. This definition of interestingness considers that novel is interesting and degrees of novelty may be labeled as degrees of interestingness.

3.4 Interestingness and Limitation of Association Rules

The research work surrounding interestingness measures seeks to move beyond the boundaries that association rules have created. The limitations of association rules are as follows:

1. Association rules do not apply to individual records. Because they partition data into n-dimensional cubes, association rules define interesting clusters and not the interestingness of individual records.

2. Association rules are static. They take a static snapshot of the data collected and do not accommodate the human experience, which is a dynamic reaction to stimuli, according to Berlyne [3].

3. Association rules are subjective. This aspect raises the relevant question of whether interestingness must always be subjective and able to incorporate apriori information.

4. Association rules do not accommodate novelty. According to Geng [9], novelty requires more research, and it is interesting that association rules have not yet accommodated this important aspect of statistics.

From Schmidhuber's work in information theory [1], interestingness measures research draws on two ideas relevant to the formulation of a principle of interestingness. First, it draws on the idea of *"too trivial or too random"* to reinforce the notion of "degree" of interestingness. Second, the definition of interestingness as a rate of change of the model is unique and enables relevant mathematical interpretations of the degree of interestingness of an observation.

As concerns anomaly detection, there is a sense that if something is new or different, then it is interesting, but these concepts are not in the field of statistics where they are embodied as novelty, deviations, anomalies, or outliers. While anomaly detection has not contributed to the definition of interestingness per se, the field has provided potentially useful tools and techniques.

3.5 Interestingness (re)Defined

The definition-nouveau of interestingness is based on the idea that humans form many dynamic models over time. Interestingness, as Davis explains, is that which challenges the established models. Previous definitions have not considered the question of how to motivate change in existing data models and classifications. To confront this problem, this definition defines distance from the model boundary as a measure of interestingness and includes the potential for irregular model boundaries. It makes three assumptions:

1. Structure: structure exists in the data.

2. Model Complexity: the model chosen for the data is sufficient for the task and has been properly constructed. It is implied by the definitions that the model must be able to distinguish a cluster of records, or class, from all others.

3. Model Fit: the model properly describes the data. The model chosen for the data must be sufficiently complex to capture the intricacies of the possible structure in the data.

As previously mentioned, the proposed definition also captures the notions of creativity and insight. Insight and creativity are perhaps a question of establishing the boundaries of existing models and then either adjusting the models, or seeking data along those boundaries that is of interest.

Having established the definition, it is important to consider approaches to defining interestingness within data sets. In order to assess the interestingness of a data point, any approach must be one that permits a means of locating the point in the feature space relative to the decision surfaces of the model. The use of binary classifiers for a multiclass approach allows one to triangulate the location of data points in feature space and discern both ambiguous points and the ambiguous in feature space; however, even in a simple system, ambiguous regions can be very complex. A method for assigning a measure of interestingness to a data point will necessitate a method to meaningfully consolidate the outputs of several classifiers. Outside of the approach mentioned, the research also considers the use of decision trees and neural networks as implementation techniques.

4. Conclusion and Future Work

In this paper, we reviewed pertaining literature and terminologies related to *interestingness*. To summarize, interestingness can be defined as a measure which challenges the established models. Analogously, unexpectedness is defined as a subjective notion of interestingness, while the process of retrieving the areas that are interesting for the understanding of the event is defined as anomaly detection. This paper surveyed various approaches to interestingness by provided highlights for the previous work performed by [7], [29], [8], [9], [2], [3] and [10]

Bourassa [3] noted that from the social and cognitive science perspective, *an interesting theory challenges audiences' beliefs, but not too much*. The relevant conclusion is that interesting is novel, and the degree of novelty indicates the degree of interestingness, defining "trivial" and "random" as two extremes of novelty.

The reviewed measures of interestingness offer diverse definitions for what interestingness is or may be. The definitions often depend on the patterns of the problems being addressed; future research work may seek to establish a correlation between subjective and objective measures which distills the common themes of existing interpretations of interestingness and synthesize a new, unifying definition that can be applied generally to all forms of data analysis.

References

1. Schmidhuber, J., *What's interesting?* 1997.
2. Chandola, V., A. Banerjee, and V. Kumar, *Anomaly detection: A survey.* ACM Computing Surveys (CSUR), 2009. **41**(3): p. 15.
3. Bourassa, M.A.J., *Interestingness: Guiding the Search for Significant Information*, in *Division of Graduate Studies of the Royal Military College of Canada*. 2011, Royal Military College of Canada.
4. Popescu, A., et al. *A remote sensing image processing framework for damage assessment in a forest fire scenario*. in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. 2012: IEEE.
5. Markou, M. and S. Singh, *Novelty detection: a review—part 1: statistical approaches.* Signal Processing, 2003. **83**(12): p. 2481-2497.
6. Dokas, P., et al. *Data mining for network intrusion detection*. in *Proc. NSF Workshop on Next Generation Data Mining*. 2002.
7. Hilderman, R.J. and H.J. Hamilton, *Knowledge discovery and interestingness measures: A survey*. 1999: Citeseer.
8. McGarry, K., *A survey of interestingness measures for knowledge discovery.* The knowledge engineering review, 2005. **20**(01): p. 39-61.
9. Geng, L. and H.J. Hamilton, *Interestingness measures for data mining: A survey.* ACM Computing Surveys (CSUR), 2006. **38**(3): p. 9.
10. Kontonasios, K.N., E. Spyropoulou, and T. De Bie, *Knowledge discovery interestingness measures based on unexpectedness.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge

Discovery, 2012.

11. Agrawal, R., T. Imieliński, and A. Swami. *Mining association rules between sets of items in large databases*. in *ACM SIGMOD Record*. 1993: ACM.
12. Agrawal, R. and R. Srikant. *Fast algorithms for mining association rules*. in *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*. 1994.
13. Piateski, G. and W. Frawley, *Knowledge discovery in databases*. 1991: MIT press.
14. Goodman, R.M. and P. Smyth, *Rule induction using information theory.* Knowledge Discovery in Databases (KDD-1991). MITPress, 1991. **199**.
15. Major, J.A. and J.J. Mangano, *Selecting among rules induced from a hurricane database.* Journal of Intelligent Information Systems, 1995. **4**(1): p. 39-52.
16. Klemettinen, M., et al. *Finding interesting rules from large sets of discovered association rules*. in *Proceedings of the third international conference on Information and knowledge management*. 1994: ACM.
17. Matheus, C.J., G. Piatetsky-Shapiro, and D. McNeill, *20 Selecting and Reporting What is Interesting: The KEFIR Application to Healthcare Data.* 1996.
18. Hamilton, H.J. and D.R. Fudger, *Estimating DBLearn's potential for knowledge discovery in databases.* Computational Intelligence, 1995. **11**(2): p. 280-296.
19. Silberschatz, A. and A. Tuzhilin. *On subjective measures of interestingness in knowledge discovery*. in *Proceedings of KDD-95: First International Conference on Knowledge Discovery and Data Mining*. 1995.
20. Kamber, M. and R. Shinghal. *Evaluating the interestingness of characteristic rules*. in *Proc. 1996 Int'l Conf. on Data Mining and Knowledge Discovery (KDD'96), Portland, Oregon*. 1996.
21. Hamilton, H., N. Shan, and W. Ziarko, *Machine learning of credible classifications.* Advanced Topics in Artificial Intelligence, 1997: p. 330-339.
22. Liu, B., W. Hsu, and S. Chen. *Using general impressions to analyze discovered classification rules*. in *Proc. of the 3rd International Conference on Knowledge Discovery and Data Mining*. 1997.
23. Gago, P. and C. Bento, *A metric for selection of the most promising rules.* Principles of Data Mining and Knowledge Discovery, 1998: p. 19-27.
24. Freitas, A., *On objective measures of rule surprisingness.* Principles of Data Mining and Knowledge Discovery, 1998: p. 1-9.
25. Gray, B. and M. Orlowska, *CCAIIA: Clustering categorical attributes into interesting association rules.* Research and Development in Knowledge Discovery and Data Mining, 1998: p. 132-143.
26. Dong, G. and J. Li, *Interestingness of discovered association rules in terms of neighborhood-based unexpectedness.* Research and Development in Knowledge Discovery and Data Mining, 1998: p. 72-86.
27. Liu, H., et al., *Efficient search of reliable exceptions.* Methodologies for Knowledge Discovery and Data Mining, 1999: p. 194-204.
28. Zhong, N., Y. Yao, and S. Ohsuga, *Peculiarity oriented multi-database mining.* Principles of Data Mining and Knowledge Discovery, 1999: p. 136-146.
29. Hilderman, R.J. and H.J. Hamilton, *Measuring the interestingness of discovered knowledge: A principled approach.* Intelligent Data Analysis, 2003. **7**(4): p. 347.

30. Lin, J., *Divergence measures based on the Shannon entropy.* Information Theory, IEEE Transactions on, 1991. **37**(1): p. 145-151.
31. Fabris, C.C. and A.A. Freitas, *Discovering surprising patterns by detecting occurrences of Simpson's paradox.* Research and Development in Intelligent Systems, 1999. **16**: p. 148-160.
32. De Bie, T., K.N. Kontonasios, and E. Spyropoulou. *A framework for mining interesting pattern sets*. in *Proceedings of the ACM SIGKDD Workshop on Useful Patterns*. 2010: ACM.
33. Jaroszewicz, S. and D.A. Simovici, *Interestingness of frequent itemsets using Bayesian networks as background knowledge*, in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2004, ACM: Seattle, WA, USA. p. 178-186.

Table 1 List of Probabilistic interestingness measure

Name	Formula
Information Gain	$log \frac{P(AB)}{\boxed{P(A)P(B)}}$
Goodman and Kruskal	$\frac{\sum_i max_j\ P(A_iB_j) + \sum_j max_i\ P(A_iB_j) - max_i\ P(A_i) max_i\ P(B_j)}{\boxed{2 - maxiP(Ai) - maxiP(Bj)}}$
Gini Index	$P(A)[P(B\|A)2 + P(\neg B\|A)2] + P(\neg A)[P(B\|\neg A)2 + P(\neg B\|\neg A)2] - P(B)2 - P(\neg B)2$
Example and Counter Example Rate	$1 - \frac{P(A\neg B)}{\boxed{P(AB)}}$
Coverage	$P(A)$
Cosine	$\frac{P(AB)}{\boxed{\sqrt{P(A)P(B)}}}$
Conviction	$P(A)P(\neg B)\boxed{P(A\neg B)}$
Confidence	$P(B\|A)$
Collective Strength	$\frac{P(AB) + P(\neg B\|\neg A)}{\boxed{P(A)P(B) + P(\neg A)P(\neg B)}}$
Certainty Factor	$\frac{P(B\|A) - P(B)}{\boxed{1 - P(B)}}$

Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.4, No.1, 2013

Added Value	$P(B \mid A) - P(B)$
Accuracy	$P(AB) + P(\neg A \neg B)$
Zhang	$\dfrac{P(AB) - P(A)P(B)}{\boxed{max(P(AB)P(\neg B), P(B)P(A \neg B))}}$
Yules Y	$\boxed{\dfrac{\sqrt{P(AB)P(\neg A \neg B)} - \sqrt{P(A \neg B)P(\neg AB)}}{\sqrt{P(AB)P(\neg A \neg B)} + \sqrt{P(A \neg B)P(\neg AB)}}}$
Yules Q	$\dfrac{P(AB)P(\neg A \neg B) - P(A \neg B)P(\neg AB)}{\boxed{P(AB)P(\neg A \neg B) + P(A \neg B)P(\neg AB)}}$
Two-Way Support	$\dfrac{P(AB) \times log2P(AB)}{\boxed{P(A)P(B)}}$
Support	$P(AB)$
Specificity	$P(\neg B \mid \neg A)$
Sebag Schoenauer	$\dfrac{P(AB)}{\boxed{P(A \neg B)}}$
Relative risk	$\dfrac{P(B \mid A)}{\boxed{P(B \mid \neg A)}}$
Recall	$P(A \mid B)$
Prevalence	$P(B)$
Piatesky Shapiro	$P(AB) - P(A)P(B)$
Odd's Ratio	$\dfrac{P(AB)P(\neg A \neg B)}{\boxed{P(A \neg B)P(\neg BA)}}$
Odd Multiplier	$P(AB)P(\neg B)\boxed{P(B)P(A \neg B)}$
Normalize Mutual Information	$\dfrac{\Sigma i \Sigma j P(AiBj) \times log2(AiBj)}{\boxed{P(Ai)P(Bj) - \Sigma i P(Ai) * log2P(Ai)}}$
Loevinger	$\dfrac{1 - P(A)P(\neg B)}{\boxed{P(A \neg B)}}$
Linear Correlation Co-efficient	$\dfrac{P(AB) - P(A)P(B)}{\boxed{\sqrt{P(A)P(B)P(\neg A)P(\neg B)}}}$

Lift	$\frac{P(B\|A)}{P(B)}$ or $\frac{P(AB)}{P(A)P(B)}$
Leverage	$P(B\|A) - P(A)P(B)$
Least Contradiction	$\frac{P(AB) - P(A\neg B)}{P(B)}$
Laplace Correction	$\frac{n(AB) + 1}{n(A) + 2}$
Klosgen	$\sqrt{P(AB)}(P(B\|A) - P(B)), \sqrt{P(AB)}max(P(B\|A) - P(B), P(A\|B) - P(A))$
JMeasure	$P(AB)log(P(B\|A)P(B)) + P(A\neg B)log(P(\neg B\|A)P(\neg B))$
Jacard	$\frac{P(AB)}{P(A) + P(B) - P(AB)}$
IWD	$((\frac{P(AB}{P(A)P(B))})^k - 1) \times P(AB)^m$