A Comprehensive Study on Risk Assessment techniques in development of Sensitive Information Management System

D. Naga Malleswari*, K. Subrahmanyam
Department of Computer Science and Engineering, KL University, Green fields, Vaddeswaram, Guntur (Dist.),
Andhra Pradesh, India.
*Corresponding author: E-Mail: nagamalleswary@kluniversity.in

ABSTRACT

Risk assessment involves a disciplined approach of maintain information to support a risk decision, which consists of identification and analysis of hazards and associated risk with the occurrence of these hazards. In this paper, we examine about the risk assessment of delicate information processing strategies in Information Systems like pharmacy, health, automobile, and educational systems. Primarily risk assessment methods are arrange into two classifications Qualitative investigation techniques and Quantitative examination strategies. We made a comprehensive study on Qualitative approaches such as Fuzzy sets and Rough sets and Quantitative approaches such as AHP, Bayesian classifier and Bayesian networks and compared the advantages and disadvantages of the above mentioned methods for given sensible data and concluded the best suitable assessment methods for the sensitive pharmacy data.

KEY WORDS: Risk, AHP, Bayesian Networks, Fuzzy, Soft computing Sensitive Information Systems.

1. INTRODUCTION

Risk management: Can be defined as identify, analyze evaluation, control, and avoidance, reduce, or eliminate the unacceptable risks. A risk is a potential event that will adversely affect the ability of a system to perform its mission should the risk event take place. A risk has two basic attributes, Probability P and Impact I, where Probability stands for the likelihood that an event will occur. A risk Ri can be defined as Ri= function (probability, impact). A common procedure to identify risk is sequential process. Each and every individual component will check and identify the risk. Using modern technology identification of risk is small issue but IT people more careful about to dependency risk that means if once we identify the risk depending upon the risk what impact is going on and if any changes are there in remaining un checking project files all are identify and reduce risks.

In real world user may not think about dependency risks. For example user is working with a product he may face an error, due to in work busy he may overlook the problem and tried for alternative paths to do the things. In this case a new error erases due to neglecting the previous error. Dependency risks are more harm full sometimes entire system works in negative directions. Robots manufacturing process dependency risks place major role. If we develop hundred percent risk free IT products automatically we provide efficient information security product (Chaitanya Krishna and Kodukula Subrahmanyam, 2015).

Now a day's maximum researchers are doing Qualitative approach because Quantitative approach is very difficult, we can measuring a numerical value for each and every phase. Qualitative approach is not suitable for health care system. The result of the qualitative approach is depend up on the researchers taken value one researchers to another researchers assumptions are different so this is not suggestible for health care system (Chaitanya Krishna and Kodukula Subrahmanyam, 2015).

2. METHODS & MATERIALS

Qualitative Methods

Fuzzy Sets: Fuzzy interpolative thinking is an essential deduction procedure for scanty fluffy principle based frameworks, where the fluffy sets showing up in the predecessors of the fluffy standards don't spread the entire data universe of talk. At the end of the day, there is an unfilled space between two nearby enrollment elements of the fluffy sets showing up in the precursors of the fluffy standards. In the event that a perception happens in the unfilled space, there is no tenet terminated and no result will be derived. As of late, some fluffy interpolative thinking strategies have been introduced for inadequate fluffy standard based frameworks. Peter Baranyi (1999), displayed a fluffy introduction strategy that can surmise the fluffy interpolative thinking result in light of the fluffy connection and the semantic connection of fluffy sets. Peter Baranyi (1999), displayed a fluffy interpolative thinking strategy that changes the α-slice based guideline insertion technique to maintain a strategic distance from irregular conclusions. Bouchon and Marsala (2000), introduced a fluffy interpolative thinking technique in light of the idea of graduality, which derives a conclusion by method for the changes of area and shape gradually. Haung (2015), introduced a fluffy interpolative thinking strategy in light of the delegate estimations of fluffy sets and exhibited their scale and move change administrators to manage fluffy interpolative thinking. Haung (2015), enhanced the technique to handle numerous fluffy standards insertion and fluffy tenet extrapolation. Yam and Koczy (1997), exhibited a direct fluffy interpolative thinking technique, which utilizes the extents of fluffy separation between the perception and principle predecessors to gather the fluffy interpolative thinking result. Bouchon and Marsala (2000),

exhibited a fluffy interpolative thinking strategy with different variable guidelines. Yam and Koczy's (1997), fluffy interpolative thinking technique in inadequate fluffy tenet based frameworks. Yam and Koczy's (1997), fluffy interpolative thinking technique does not generally prompt arched conclusions. Peter Baranyi (1999), looked at the adjusted α-cuts-based fluffy interjection strategy and Yam and Koczy's (1997), fluffy interpolative thinking technique and made a far reaching investigation. Tao Wang and Yan Shi (2011), introduced another fluffy interpolative thinking strategy to deduce a conclusion taking into account the similitudes of fluffy sets. Yam and Koczy (1997), exhibited a fluffy interpolative thinking technique with capacity space representation of enrollment capacities. Yam and Koczy (1997), displayed a technique for speaking to participation capacities as focuses in high-dimensional spaces for fluffy interjection and fluffy extrapolation.

Rough Sets: There are numerous elements that persuade the incorporation of a dimensionality lessening (DR) venture in an assortment of critical thinking frameworks. Numerous application issues process information as an accumulation of genuine esteemed vectors (for instance, content characterization, bookmark classification). In the event that these vectors show a high dimensionality, then handling gets to be infeasible. Along these lines, it is regularly helpful, and now and again important to diminish the information dimensionality to a more reasonable size with as meager data misfortune as could be allowed. Some of the time, high-dimensional complex marvels can be administered by essentially less, straightforward variables. The procedure of dimensionality diminishment here will go about as an instrument for displaying these marvels, enhancing their clarity. There is frequently a lot of excess or deluding data display; this should be evacuated before any further preparing can be completed. For instance, the issue of getting grouping rules from huge datasets regularly advantages from an information lessening pre-preparing step. Not just does this diminish the time required to perform incitement, yet it makes the subsequent standards more understandable and can build the subsequent grouping accuracy. Though semantics-wrecking dimensionality lessening procedures irreversibly change information, semantics-saving DR strategies (alluded to as characteristic determination) endeavor to hold the significance of the first trait set. The primary point of credit choice is to decide a negligible trait subset from an issue area while holding a suitably high precision in speaking to the first qualities. There are frequently numerous properties included, and combinatorially extensive quantities of trait blends, to choose from. Note that the quantity of property subset mixes with characteristics from an accumulation of aggregate qualities is. It may be normal that the consideration of an expanding number of characteristics would improve the probability of including enough data to recognize classes. Shockingly, this is not genuine if the extent of the preparation dataset does not likewise increment quickly with each extra quality included. This is the supposed condemnation of dimensionality. A high-dimensional dataset builds the odds that an information mining calculation will discover spurious examples that are not legitimate by and large. Most strategies utilize some level of decrease keeping in mind the end goal to adapt to a lot of information, so a productive and viable lessening strategy is required. A strategy that can diminish dimensionality utilizing data contained inside of the information set and that jelly the significance of the characteristics (i.e., semantics-safeguarding) is unmistakably attractive. Harsh set hypothesis (RST) can be utilized all things considered an instrument to find information conditions and to diminish the quantity of traits contained in a dataset utilizing the information alone, requiring no extra information.

Quantitative Methods:

Analytical Heirarchy Process: In the logical progression process (AHP), pairwise examination frameworks (i.e., inclination relations) are utilized to express leaders' inclinations. Since the presence of the AHP, the AHP-collective choice making (GDM) issue has been a prominent subject. By and large, AHP-GDM contemplates have been introduced in a neighborhood setting (a remarkable model). For AHP-GDM in a neighborhood setting, each AHP chief gives an individual pairwise correlation lattice, and after that an AHP-GDM model is connected to get an aggregate need vector from the individual pairwise examination matrices (Yucheng Dong, 2015). The execution of the AHP-GDM in a neighborhood connection includes three procedures:

- Prioritization process: The prioritization process alludes to the determination of individual need vectors from the semantic pairwise examination lattices (LPCMs). The legitimacy of the determined individual need vector depends on two elements: the determination of numerical scale and the outline of prioritization method (Yucheng Dong, 2015).
- AHP choice model: The AHP choice model is applied to acquire an aggregate need vector to request options. In existing AHP choice models, the conglomeration of individual judgments (AIJ) and the collection of individual needs (AIP) are generally used (Yucheng Dong, 2015).
- AHP accord demonstrate: The agreement model helps chiefs come to a consensus (Yucheng Dong, 2015).

Naive Bayesian Classifier: Innocent Bayesian classifier (NBC) or Naïve Bayes (NB) is a rearranged type of Bayes' decide that expect autonomy of the perceptions. Some exploration result, showed that NBC has aggressive execution in examination with other learning calculations if the ordinary conveyance suspicion holds (James, 2001). Picture characterization utilizing ordinary Bayesian classifier is a directed strategy in view of earlier probabilities. Earlier likelihood determines disarrays among classes that are not well divisible and is along these lines successful in

enhancing characterization accuracy (James, 2001). Bayesian classifier is for the most part utilized with the presumption that earlier probabilities are equivalent, as dependable earlier probabilities are not effectively accessible. Notwithstanding, with the utilization of equivalent earlier probabilities, the execution of the classifier is not ideal, which is clear from the misclassified pixels essentially because of unearthly perplexity between classes, aside from the expanded registering and testing requirements (James, 2001). Prior works utilized earlier probabilities taking into account earlier year crop insights, geological information, height information, and spatial qualities indicated through a Markov irregular field model at reference determination, enhancing the general precision and kappa coefficient (James, 2001). In this manner, it is attractive to acquire dependable earlier probabilities for each otherworldly class and utilize them to characterize the pixels that are liable to misclassify, despite the fact that they are hard to decide inside of the same time period and for the same unearthly classes (James, 2001).

Bayesian Networks: The fast development of the Internet makes it simple to gather information on an extensive scale (Zhiqiang Yang, 2006). Information is for the most part put away by various elements, extending from people to little organizations to government offices. This information incorporates delicate information that, if utilized despicably, can hurt information subjects, information proprietors, information clients, or other applicable parties (Zhiqiang Yang, 2006). Worry about the proprietorship, control, protection, and precision of such information has turned into a top need in specialized, scholarly, business, and political circles. At times, regulations and shopper backfire additionally deny distinctive associations from imparting their information to each other. Such regulations incorporate HIPAA and the European security mandates. As a sample, consider a situation in which an exploration focus keeps up a DNA database around a vast arrangement of individuals, while a doctor's facility stores and keeps up the history records of those individuals' therapeutic conclusions. The examination focus needs to investigate connections between DNA successions and particular sicknesses. Because of security concerns and protection regulations, the doctor's facility can't give any data about individual therapeutic records to the examination focus. Information mining customarily requires all information to be assembled into a focal site where particular mining calculations can be connected on the joint information. This model works in numerous information mining settings. In any case, unmistakably this is undesirable from a security point of view. Disseminated information mining evacuates the prerequisite of conveying every single crude data to a focal site, however this has typically been spurred by reasons of proficiency and arrangements don't as a matter of course give protection. Interestingly, security saving information mining arrangements, including our own, give information mining calculations that register or surmised the yield of a specific calculation connected to the joint information, while ensuring other data about the information. Some security saving information mining arrangements can likewise be utilized to make changed, publishable variants of the data information sets (Zhiqiang Yang, 2006). Bayesian systems are an intense information mining device. A Bayesian system comprises of two sections: the system structure and the system parameters. Bayesian systems can be utilized for some errands, for example, theory testing and robotized experimental discovery (Zhiqiang Yang, 2006).
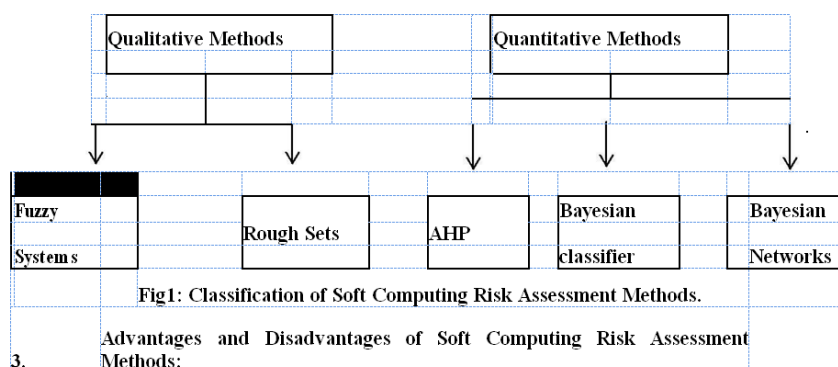
Soft computing Risk:

Fig1: Classification of Soft Computing Risk Assessment Methods.

3. Advantages and Disadvantages of Soft Computing Risk Assessment Methods:

Fig.1. Assessment Methods

Advantages and Disadvantages of Soft Computing Risk Assessment Methods:
Fuzzy systems:
Advantages: a) Flexible, natural information base design, Control and Supervisions peak the same language. b) Convenient client Interface, less demanding end-client translation when the final client is not a control Engineer. c) Easy calculation, widely available tool compartments and devoted coordinated Circuits. d) Validation, Consistency, redundancy and culmination can be checked in standard bases that could accelerate computerized Learning and Enhance client Interpretability.

Disadvantages: a) Experimental Manual tuning in vast scale mechanical applications. Tedious returning regardless of the fact that connected to a comparative plant in other location. b) Intuitive fluffy PID-like outline does not plainly beat very much tuned ordinary controllers. c) Dimensionality, Cartesian result of allotments is the most utilized

method for setting up forerunners in multi- dimensional models is an exceptionally wasteful and memory concentrated setting for generally Functions.

Rough Sets:

Advantages: a) Identifying connections that would not be found while utilizing measurable Methods. b) Allowing both subjective and quantitative data. c) Finding insignificant arrangements of information reduction. d) Evaluating centrality of information. e) Generating arrangements of choice tenets from data.

Disadvantages: a) Inefficiency of unpleasant set techniques. b) Identifying the nonessential qualities, which restrains the suitability of the customary unpleasant set model in information mining applications. c) Relational database frameworks and a great deal of computational escalated operations are performed in level records as opposed to using the superior database set operations.

AHP:

Advantages: a) AHP can mull over the relative needs of elements or options and speaks to the best alternative. b) AHP gives a straight forward and extremely adaptable model for a given Problem. c) AHP gives a simple relevant choice settling on approach that help the chief to unequivocally choose the judgments. d) Either goal or subjective contemplations or either quantitative or subjective data assume an imperative part amid the choice process.

Disadvantages: a) There is not generally an answer for the direct equations. b) The computational prerequisite is colossal not withstanding for a little Problem. c) AHP permits only triangular fluffy numbers to be used. d) AHP depends on both likelihood and probability measures.

Bayesian Classifier:

Advantages: a) Provides a characteristic and principled method for joining earlier data with Data. b) Provides derivations that are restrictive on the information and are precise, without dependence on asymptotic Approximation. c) Provides an advantageous setting for an extensive variety of models, for example, various leveled models and missing information problems.

Disadvantages: a) It doesn't let you know how to choose a prior. b) It can deliver back conveyances that are vigorously affected by the Priors. c) It regularly accompanies a high computational expense, particularly in models with a substantial number of parameters.

Bayesian Networks:

Advantages: a) Straight forward representation of causal connections between framework variables. b) Utilize an assortment of information, Representation of uncertainty. c) Visual choice bolster instrument, Can deal with missing observations.

Disadvantages: a) Troublesome achieving assertion on the BN structure with experts. b) Troublesome characterizing the CPTs with master conclusion furthermore No input loops. c) Continuous information representation furthermore Spatial and temporal.

3. RESULTS

Quantative Methods:

Rough sets: They are applicable in the developing the less demanding interface when the end user is not an control engineer.

Fuzzy sets: This set of the quantitative method is applicable while identifying the connections between the data and evaluating the centrality of the system.

Qualitative Method:

AHP: This is used to identify the best possible way of solving the problem and giving the direct answer to the question.

Bayesian Classifier: This method is used in integrating the earlier data with the present data and derivation is provided in the précised manner without depending upon the notations.

Bayesian Networks: This method is useful in representing the casual connections of the network variables analysing the missing observations using the visual choice bolster instrument.

4. CONCLUSION

In this paper, we talk about the danger Assessment of Soft Computing Methods in data frameworks like Fuzzy sets, Rough sets, AHP, Bayesian Classifier and Bayesian Networks. Further study will be on Hybrid model for Risk Assessment in data frameworks on different fields. It is created by Integrating two or all the more existing models like Rough-Bayesian systems, Fuzzy-harsh sets, Fuzzy-AHP, Gray-Hierarchy model, GA-based neural system approach.

REFERENCES

Bouchon-Meunier B, Marsala M.C, Interpolative reasoning based on graduality, IEEE, 5, 2000, 483-487.

Chaitanya Krishna B, Kodukula Subrahmanyam, A Dependency analysis for Information Security and Risk Management, International Journal of Security and Its Applications, 9 (8), 2015, 205-210.

Chaitanya Krishna B, Kodukula Subrahmanyam, A Novel Approach for Information Security and Risk Management in Health Care Systems, International Journal of Bio-Science and Bio-Technology, 7 (4), 2015, 11-16.

Haung, Intelligent Robotics and Applications, 8th International Conference, ICIRA 2015, Portsmouth, UK, August 24-27, 2015.

James N.K, Liu, Bavy N.L, Li and Tharam S, Dillon, An Improved Naïve Bayesian Classifier Technique Coupled With a Novel Input Solution Method, IEEE Transactions On Systems, Man, And Cybernetics—Part C, Applications And Reviews, 31 (2), 2001.

Peter Baranyi, Domonkos Tikk, Yeung Yam, Laszlo T. Koczy, Investigation of a new α-cut based fuzzy interpolation method, CUHK-MAE-9906, Dept. Mechanical and Automation Eng, The Chinese University of Hong Kong, Hong Kong, 1999.

Tao Wang and Yan Shi, Fuzzy Interpolative Reasoning and Algorithms for the Sparse Fuzzy Rule, International journal of innovative computing information and control, ICICI, 7, 2011, 1689-1697.

Yam Y and Koczy L.T, Representing membership functions as points in high dimensional spaces for fuzzy interpolation and extrapolation, Technical Report CUHK-MAE-97-03, Dept of Mechanical and automation Eng, The Chinese Univ. of Hong Kong, 1997.

Yucheng Dong, Zhi-Ping Fan, and Shui Yu, Consensus Building in a Local Context for the AHP-GDM With the Individual Numerical Scale and Prioritization Method, IEEE Transactions On Fuzzy Systems, 23 (2), 2015.

Yu-Chuan Chang, Shyi-Ming Chen, and Churn-Jung Liau, Fuzzy Interpolative Reasoning for Sparse Fuzzy-Rule-Based Systems Based on the Areas of Fuzzy Sets, IEEE Transactions On Fuzzy Systems, 16 (5), 2008.

Zhiqiang Yang and Rebecca N, Wright, Privacy-Preserving Computation of Bayesian Networks on Vertically Partitioned Data, IEEE Transactions On Knowledge And Data Engineering, 18 (9), 2006.