ExVis: Explainable Visual Decision Support System for Risk Management

Md Rafiqul Islam*, Jiaming Zhang*, Md. Hamjajul Ashmafee†, Imran Razzak ‡,
Jianlong Zhou§, Xianzhi Wang §, Guandong Xu *

Advanced Analytics Institute (AAi), University of Technology Sydney, NSW, Australia
†*Department of CSE, Islamic University of Technology (IUT), Dhaka, Bangladesh*
‡*School of Information Technology, Deakin University, Geelong, Australia*
§*School of Computer Science, University of Technology Sydney (UTS), Australia*

{MdRafiqul.Islam-1, Jiaming.Zhang-3}@student.uts.edu.au, ashmafee@iut-dhaka.edu, Imran.Razzak@deakin.edu.au
{Xianzhi.Wang, Jianlong.Zhou, Guandong.Xu}@uts.edu.au

Abstract—Managing and communicating risk has become a crucial task in today's economy and society, where data visualization can play a vital role. Existing researchers have applied various methods to assess the risks and have shown traditional ways of visualization such as maps, charts, diagrams, etc. However, designs, developments, evaluations, and lack of proper explanations have decreased their reliability in decision-making. In this regard, we propose a solution named "Explainable visual decision support system (*ExVis*)" to present all the relevant information concisely and provide a visual recommendation for risk management. *ExVis* combines the state-of-the-art interactive visualizations blending with the Bayesian network (BN) model. Our approach is used to analyze insurance data as a case study. In this study, we have three main implications: (i) insurance data are potentially useful for understanding business risk; (ii) a visual explanation dashboard enhances stakeholder decision-making; and (iii) some critical insights into the underwriting of policyholders have been deduced. Finally, we evaluate the usefulness of our visual interactive system (VIS) by a user study.

Index Terms—Decision support system, explainable visualization, multiple coordinated views, risk management.

I. INTRODUCTION

The financial industry, such as the life insurance industry, is becoming complex daily, where effective communications between risk experts and decision-makers play a vital role [1, 6, 14]. Visualization is one of the efficient ways to support such effective risk-related communication [4, 9]. Nowadays, many types of business strategies, mapping approaches, and visual metaphors are employed in various business purposes [2]. However, risk visualization is still rare in the business community because the risk is both challenging to visualize and hard to manage.

A fundamental business aspect of the business market is assessing the risk connected with each individual [11]. For example, policyholders with mental health illness (MHI) cause a financial imbalance and affect their professional and personal lives [6, 7]. The existing studies show that the qualitative and quantitative aspects have been dominated to analyze and manage the risks. Tables, charts, maps, and formulas are standard tools in daily communication between experts and managers regarding financial risks and how to handle them [2]. However, it is challenging for decision-makers to identify and understand the most significant risks and implement adequate remedies when considering policyholders' negative experiences in seeking insurance benefits. Therefore, there is no doubt that insurance companies make a loss, where an interactive solution is required to justify the claiming insurance benefits with proper explanation.

Visualization systems are used to analyze and visualize data using machine learning (ML), deep learning (DL), statistical models, etc. [5, 8, 10]. For example, Varga [16] demonstrated how visual analytics (VA) could help financial services thrive. MHIVis [7] is a new VIS to provide a visual summary of mental health disorder data. However, some VIS still incomprehensible and less appealing in many situations when it comes to decision-making [12, 13]. As a result, the concerns of presenting the data and why a particular outcome is recommended remain unresolved [3, 18].

We design an interactive explainable visual decision support system (*ExVis*) for risk management that combines a Bayesian network (BN) model with visual recommendations to respond to these challenges. Fig. 1 illustrates how *ExVis* enables IMs to assess policyholder history to recommend future claim records. The dashboard evaluates 10 years of data from $31,800$ policyholders to provide a coordinated view of accidents and diseases across Australia. Prior policyholder behaviors such as alcohol, travel, smoking, and diabetes were examined, as shown in Table I. As a result, the system analyses data to produce a likelihood score from multiple perspectives. In summary, the following **key contributions** are made by our work as follows:

- We propose an observational research focusing on insurance data potentially helpful in understanding business risk.
- We develop an explainable visual decision support system named *ExVis* which allows decision-makers to justify policyholders claim benefits.
- We provide some critical insights into the underwriting of policyholders for insurance managers to help make decisions.

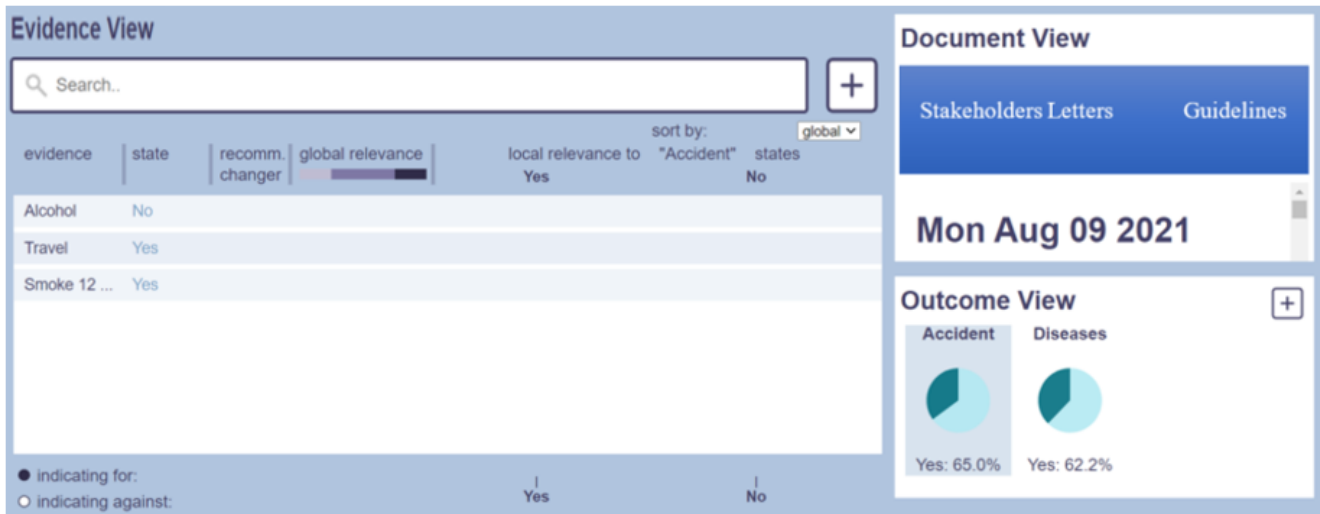The following is the outline for this paper. In Section II,

Fig. 1: Exploring multiple views to explain visual decision support system

TABLE I: Characteristics of respondents

Risk Category	Risk Factor
Diseases	Ashma
	Diabetes
	Rel Diabetes
	Cancer
Accident	Travel
	Alcohol
	Smoking 12 months
	Smoking 5 Years

we briefly discuss the methodology followed by the detail our visual approach named "*ExVis*" in section III. Finally, we describe our user study in IV, and draw conclusions in section V.

II. METHODOLOGY

This section discusses how *ExVis* is used as a designed dashboard for examining visual explanation for risk management and the details of data gathering and processing.

A. Data collection and processing

We collected three different types of data from Australian insurance companies, including (i) questionnaire data- the questionnaire was extensive and detailed, containing $31,800$ policyholders and 834 questions ranging from personal, medical, family history, occupational details, lifestyle, etc. (ii) demographic data- the demographic data of each user has various attributes including Insurance ID, Gender, Age, Occupation, and Postcode; and iii) policyholders' claims - the claims dataset consists of $27,458$ claim records containing 37 attributes that occurred throughout the country.

B. Study Requirements and Design

In the insurance domain, decision-making is generated based on insurance guidelines, but the knowledge and expertise of insurance managers play a crucial role. The guidelines consider policyholder-specific data, such as Insurance ID,

Gender, Age, Occupation, and Postcode. Such data is available from various sources but primarily present in unstructured and unsorted forms, challenging decision-making. IMs analyze available information and integrate it with their experience to make the decision. This process automatically filters out all appropriate information entities likely to be impacted by the expected outcome. Therefore, considering the proposed design and functional criteria of *ExVis*, our approach must include representations of: **R1-Policy-specific data**: IMs must be aware of facts that support and contradict the recommendation, as well as its potential alternatives. **R2-Policy Guidelines**: IMs must be mindful of the diseases and accidental evidence used for generating a recommendation. **R3-Reasoning model**: IMs must comprehend the underlying reasoning process for improving the acceptance and trustworthiness of a recommendation. **R4-Decision**: it is necessary to demonstrate the accuracy of the computed recommendation. On the other hand, it must convey its uncertainty.

C. Bayesian Network Model

Bayesian networks, a probabilistic graphical model, are being developed to assist people in making challenging decisions about uncertainty management. These systems are at the confluence of artificial intelligence, machine learning, and statistics. BN, developed based on Bayes' theorem, proposed by Thomas Bayes, is the basis for the Bayesian network. It calculates the posterior probability for a target of interest given a collection of input values usually referred to as the findings or evidence. It relates to inference on the target value. Based on the directed acyclic graph (DAG), which comprises a collection of nodes and directed arcs, statistical connections may represent directed arcs. Nodes indicate system variables, and an arc represents a cause-and-effect connection or dependencies between those variables in this graph, where node confidence values are calculated by combining conditional probabilities [15]. BN is resilient and compact as it can make
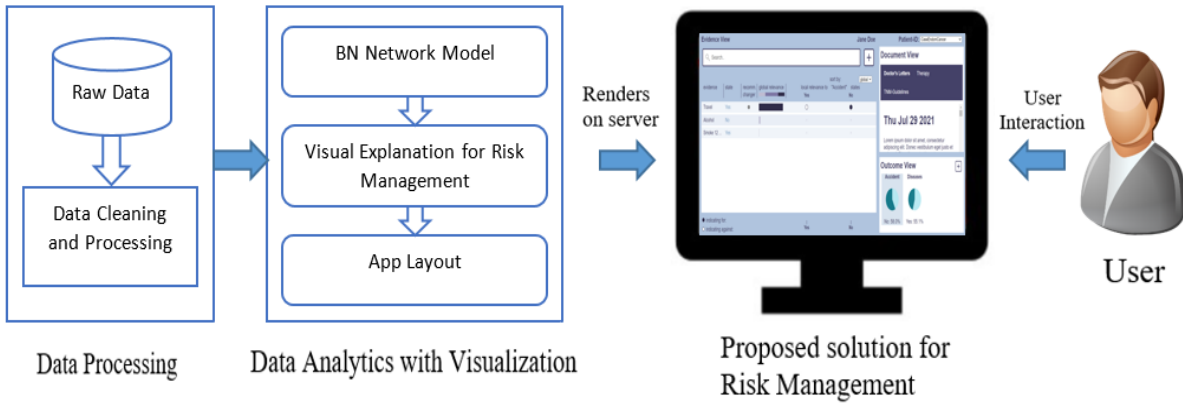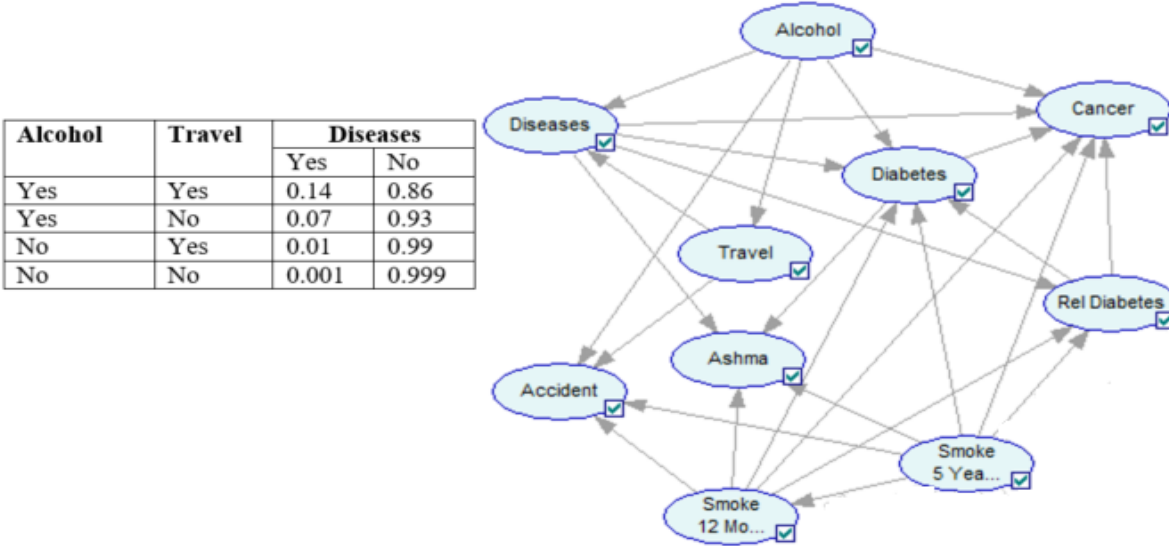
Fig. 2: The framework of methodology

Alcohol	Travel	Diseases	
		Yes	No
Yes	Yes	0.14	0.86
Yes	No	0.07	0.93
No	Yes	0.01	0.99
No	No	0.001	0.999

Fig. 3: The structure of Bayesian Network Model

TABLE II: Conditional probability table (CPT) for Bayesian network (BN)

Risk factors	Ashma	Diabetes	Rel-Diabetes	Cancer	Travel	Alcohol	Smoking 12 Months	Smoking 5 Years
Ashma	N	N	N	N	N	N	N	N
Diabetes	Y	N	N	Y	N	N	N	N
Rel-Diabetes	N	Y	N	Y	N	N	N	N
Cancer	N	N	N	N	N	N	N	N
Travel	N	N	N	N	N	N	N	N
Alcohol	N	Y	N	Y	Y	N	N	N
Smoking 12 Months	Y	N	N	Y	N	N	N	N
Smoking 5 Years	Y	Y	Y	Y	N	N	Y	N

sound predictions even when some factors are unavailable to its inherent nodal dependencies [17].

One of the critical concepts of BN is conditional independence which defines that given the values of its parents (PT), a variable X is conditionally independent of its non-descendants (ND) as follows:

$$P(X|PT(X), ND(X)) = P(X|PT(X)) \qquad (1)$$

Another essential concept is the conditional probability table

(CPT) which tabulates X's distribution for potential value assignments to its parents for each node X [17]. Thus, in this study, CPTs were developed by experts' group discussions as shown in Table II. Their expertise and professions have been elaborated on earlier in this paper.

D. Exploring Visual Explanation by Bayesian Networks

According to Muller et al. [13], global and local relevance is distinguished and determined for computing the cost of omitting each piece of evidence. First, Kullback-Leibler divergence

(KLD) is used as a cost function to assess the dissimilarity between the target variable's probability distributions before (P) and after (Q) omission of evidence. The KL divergence is computed as an integral for continuous random variable distributions P and Q:

$$KLD(P||Q) = \int p(x)log(\frac{p(x)}{q(x)})dx \tag{2}$$

Since we deal with discrete variables rather than continuous ones in our proposed system, the KLD is computed as the sum of P and Q, where P and Q denote the probability distributions of discrete random variables.

$$KLD(P||Q) = \sum p(x)log(\frac{p(x)}{q(x)})dx \tag{3}$$

However, the KL divergence is not symmetrical because it doesn't follow the commutative law of two variables. Thus, the Jensen-Shannon divergence (JSD) measures the difference (or resemblance) between two probability distributions, which resolves the prior problem. It calculates a symmetrical normalized score using the KLD. This implies that P's divergence from Q is the same as Q's divergence from P:

$$JS(P||Q) = JS(Q||P) \tag{4}$$

We used Jensen-Shannon divergence (JSD) to quantify dissimilarity instead of cross-entropy based on the presence and absence of any evidence. The following formula can be used to compute the JS divergence:

$$JSD(P||Q) = \sqrt{\frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M)} \tag{5}$$

where M is calculated as:

$$\text{and } M = \frac{1}{2}(P + Q) \tag{6}$$

JSD delivers a global level of relevance but none at the local level. As a result, we calculate the signed pre-post probability differences during the sensitivity analysis and then observe the shifts of its sensitivity.

III. Visual Explanation for Risk Management

ExVis is implemented as a visual interactive explanation system as shown in Figure 2. We introduce our approach using a small BN for insurance risk management as shown in Figure 3. It consists of 10 nodes, such as alcohol, cancer, diabetes, rel-diabetes, travel, ashma, smoke 5 years, smoke 12 months, accident and diseases as shown in Table I, and 22 relations. It computes a recommendation for the most rightful claim, "Accident and diseases." The findings of our system for visualizing the RM information of policyholders are discussed as follows:

Evidence view: The evidence view shows all evidence items organized according to their level of relevance for

TABLE III: Open-ended questions for user study

SN	Questions
1	I found that *ExVis* is more interactive for exploring visual explanation.
2	*ExVis* is easy and enjoyable to use.
3	*ExVis* was worthwhile to explore explainable visual insights.
4	*ExVis* allowed me to extract accidental and diseases information.
5	I feel confident to use *ExVis* for discovering information from alcohol, smoke, cancer and so on.

the computed outcome, e.g., a "Travel, Alcohol, Smoke 12 Months" recommendation.

Document view: The document view provides accident and disease-relevant information through the policy guidelines.

Outcome view: The computed probability distributions and a configurable set of accidental and disease scores are displayed in the outcome view to show the recommendations.

Network view: The network view is exposed by selecting evidence from the evidence view. This view is essential when IMs analyze the influence of new evidence and when they disagree with the evidence's outcome.

IV. User study

To investigate the potential usability and effectiveness of *ExVis*, we invited 5 users to participate in our user study whose age range is around 20+. The participants are primarily students and academic staff, where there are 3 males and 2 females who have expertise in visualization. As shown in Table III, we suggested that participants use our system and respond to some open-ended questions about the design. The users actively participated and provided their comprehensive judgment. The users' complete judgement was that the system would effectively understand, investigate, and recommend. Finally, we observed that *ExVis* could help the analysts to gain attention for specific risks management.

V. Conclusion

We presented a decision-making workflow of an insurance claim, where we provided interactive solutions assisting in understanding and justifying the process of accidental and disease claim recommendation computation within *ExVis*. The interactive *ExVis* help the decision-makers to analyze policyholders' claims and allows for building trust in the computed recommendation and its application within policy guidelines as an additional objective decision-making opinion. The tool, *ExVis* can be used as a practical application to analyze insurance claims. Experiment analysis showed the effectiveness of *ExVis* in decision making. In essence, we provide the following **key insights:** (i) *ExVis* facilitates IM to make a decision. (ii) IMs may connect the BN's rationale to their mental model and reasoning. (iii) evidence items represent the critical influences for the recommendation. In the future, we'd like to include more input data to display various factors for exploring information visualization. In addition, we intend to test our system with an eye tracker to see what benefits multiple coordinated views can provide.

REFERENCES

[1] Coussement, K., De Bock, K.W., 2013. Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning. Journal of Business Research 66, 1629–1636.

[2] Eppler, M.J., Aeschimann, M., 2009. A systematic framework for risk visualization in risk management and communication. Risk Management 11, 67–89.

[3] Han, Y., Rozga, A., Stasko, J., Abowd, G.D., 2013. Visual exploration of common behaviors for developmental health. Visual Analytics in Healthcare .

[4] Huang, M.L., Liang, J., Nguyen, Q.V., 2009. A visualization approach for frauds detection in financial market, in: 2009 13th International Conference Information Visualisation, IEEE. pp. 197–202.

[5] Islam, M.R., Akter, S., Ratan, M.R., Kamal, A.R.M., Xu, G., 2021a. Deep visual analytics (dva): Applications, challenges and future directions. Human-Centric Intelligent Systems 1, 3–17.

[6] Islam, M.R., Liu, S., Biddle, R., Razzak, I., Wang, X., Tilocca, P., Xu, G., 2021b. Discovering dynamic adverse behavior of policyholders in the life insurance industry. Technological Forecasting and Social Change 163, 120486.

[7] Islam, M.R., Liu, S., Razzak, I., Kabir, M.A., Wang, X., Tilocca, P., Xu, G., 2020a. Mhivis: Visual analytics for exploring mental illness of policyholders in life insurance industry, in: 2020 7th International Conference on Behavioural and Social Computing (BESC), IEEE. pp. 1–4.

[8] Islam, M.R., Liu, S., Wang, X., Xu, G., 2020b. Deep learning for misinformation detection on online social networks: a survey and new perspectives. Social Network Analysis and Mining 10, 1–20.

[9] Islam, M.R., Razzak, I., Wang, X., Tilocca, P., Xu, G., 2021c. Ucbvis: understanding customer behavior sequences with visual interactive system, in: 2021 International Joint Conference on Neural Networks (IJCNN), IEEE. pp. 1–8.

[10] Kim, D.H., Hoque, E., Agrawala, M., 2020. Answering questions about charts and generating visual explanations, in: Proceedings of the 2020 CHI conference on human factors in computing systems, pp. 1–13.

[11] Leite, R.A., Gschwandtner, T., Miksch, S., Kriglstein, S., Pohl, M., Gstrein, E., Kuntner, J., 2017. Eva: Visual analytics to identify fraudulent events. IEEE transactions on visualization and computer graphics 24, 330–339.

[12] Müller, J., Cypko, M., Oeser, A., Stoehr, M., Zebralla, V., Schreiber, S., Wiegand, S., Dietz, A., Oeltze-Jafra, S., 2021. Visual assistance in clinical decision support .

[13] Müller, J., Stoehr, M., Oeser, A., Gaebel, J., Streit, M., Dietz, A., Oeltze-Jafra, S., 2020. A visual approach to explainable computerized clinical decision support. Computers & Graphics 91, 1–11.

[14] Niu, Z., Cheng, D., Zhang, L., Zhang, J., 2018. Visual analytics for networked-guarantee loans risk management, in: 2018 IEEE Pacific Visualization Symposium (PacificVis), IEEE. pp. 160–169.

[15] Sajid, Z., Khan, F., Zhang, Y., 2017. Integration of interpretive structural modelling with bayesian network for biodiesel performance analysis. Renewable Energy 107, 194–203.

[16] Varga, D., 2017. Fintech, the new era of financial services. Vezetéstudomány-Budapest Management Review 48, 22–32.

[17] Yap, G.E., Tan, A.H., Pang, H.H., 2008. Explaining inferences in bayesian networks. Applied Intelligence 29, 263–278.

[18] Zhang, Y., Chen, X., 2018. Explainable recommendation: A survey and new perspectives. arXiv preprint arXiv:1804.11192 .