ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/359132879

Causal Artificial Intelligence for High-Stakes Decisions: The Design and Development of a Causal Machine Learning Model

Article *in* IEEE Access · January 2022

DOI: 10.1109/ACCESS.2022.3155118

CITATIONS
10

READS
514

3 authors:

Bukhoree Sahoh
Walailak University
22 PUBLICATIONS **150** CITATIONS
SEE PROFILE

Kanjana Haruehansapong
Walailak University
8 PUBLICATIONS **37** CITATIONS
SEE PROFILE

Mallika Kliangkhlao
Walailak University
16 PUBLICATIONS **74** CITATIONS
SEE PROFILE

All content following this page was uploaded by Bukhoree Sahoh on 16 March 2022.

The user has requested enhancement of the downloaded file.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 12, 2022, accepted February 20, 2022. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.3155118

Causal Artificial Intelligence for High-Stakes Decisions: The Design and Development of a Causal Machine Learning Model

BUKHOREE SAHOH[1,2], KANJANA HARUEHANSAPONG[1], AND MALLIKA KLIANGKHLAO[3]

[1]School of Informatics, Walailak University, Tha Sala, Nakhon Si Thammarat 80160, Thailand
[2]Informatic Innovation Center of Excellence (IICE), Walailak University, Tha Sala, Nakhon Si Thammarat 80160, Thailand
[3]Department of Computer Engineering, Prince of Songkla University, Hat Yai, Songkhla 90112, Thailand

Corresponding author: Bukhoree Sahoh (bukhoree.sa@wu.ac.th)

ABSTRACT A high-stakes decision requires deep thought to understand the complex factors that stop a situation from becoming worse. Such decisions are carried out under high pressure, with a lack of information, and in limited time. This research applies Causal Artificial Intelligence to high-stakes decisions, aiming to encode causal assumptions based on human-like intelligence, and thereby produce interpretable and argumentative knowledge. We develop a Causal Bayesian Networks model based on causal science using *d*-separation and *do*-operations to discover the causal graph aligned with cognitive understanding. Causal odd ratios are used to measure the causal assumptions integrated with the real-world data to prove the proposed causal model compatibility. Causal effect relationships in the model are verified based on causal P-values and causal confident intervals and approved less than 1% by random chance. It shows that the causal model can encode cognitive understanding as precise, robust relationships. The concept of model design allows software agents to imitate human intelligence by inferring potential knowledge and be employed in high-stakes decision applications.

INDEX TERMS Artificial intelligence, counterfactuals, causal science, do-calculus, causal inference, cognitive computing.

I. INTRODUCTION

Critical events are unexpected situations that severely affect citizens (e.g., by causing serious injury or death), infrastructure (e.g., via transportation damage or communications failure), and government (e.g., with economic crises or financial loss). These situations lead to high pressure and life-and-death trade-offs where decision-makers must make crucial choices that may impact the daily life of millions of citizens. One constraint on making such decisions is the numerous low-probability, high-consequence situations that may arise due to uncertain and complicated factors. In addition, high-stakes decisions are limited by time and knowledge, even as they must protect against the severe consequences of failure. A knowledge discovery-based approach is required for such high-stakes management, to provide the right knowledge to the right users at the right time.

Event explanation based on Causal Artificial Intelligence (Causal AI) may interchangeably use eXplainable Artificial Intelligence (XAI) in critical event management. It is a well-known concept that is driven by observational evidence to explore knowledge for decision-makers [1], [2]. Machine Learning (ML) is a specific class of algorithms that can provide causality of Causal AI that has become an essential ingredient to serve the knowledge discovery-based approach for event explanation [3].

Ofli *et al.* [4] and Kumar *et al.* [5] employed ML-based deep neural networks using real-time evidence for detecting and explaining high-stakes events with high-performance accuracy. Formosa [6] proposed an approach for traffic conflicts using proactive safety management strategies, while Anbarasan *et al.* [7] introduced a technique for high-stakes events during flood disasters. Both support high-performance accuracy for better decision-making, but current deep learning focuses on detection and explanation performance rather than on supporting high-stakes decisions. Deep learning is notably a ''black box'', discovering events by estimating enormous sets of parameters with complicated representations. It does not provide fundamental knowledge to interpret critical events as human-like arguments, or explanations of

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Shen.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

why and how critical events happen. These missing features make it unsuitable for making high-stakes decisions.

Rudin [8] has strongly argued that the black-box model cannot offer a human-like interpretation of knowledge for the high-stakes decision process. For example, early deep learning-based black-box models reached a performance of over 95% but could not handle simple questions such as "*Why was this output predicted?*", and "*Why were other solutions not predicted?*". Only knowledge based on human-like interpretation can answer these kinds of questions and plays a key role in helping authorities understand and react to critical events. High-stakes decisions in critical event management require a new paradigm of machine learning that goes beyond general event explanation towards cognitive event interpretation.

Causal AI lets machine learning describe the cognitive reasons for predicted output based on human-like interpretations [9]. It aims to produce reasons for *"Why"* and *"How"* events happen given current evidence regardless of outcomes, and so synthesizes plausible arguments and interpretations that decision-makers can utilize. Critical event interpretation should take advantage of Causal AI-based machine learning to produce practical knowledge for high-stakes decisions. This needs causal knowledge produced by human-like intelligent agents, which will help interpret the events that may critically influence the future.

The main contributions of this research are:

- A fundamental interpretation principle based on Causal AI for high-stakes decisions;
- Causal AI-based machine learning for event interpretation-based-high-stakes decisions;
- Proof that Causal AI-based machine learning can encode high-stakes knowledge, which converges towards human-like intelligence.

The current limitations of machine learning-based high-stakes decisions are examined in section II; background on causal science based on human-like intelligence is given in section III; section IV investigates causal encoding for high-stakes decisions and its outstanding properties; section V presents a case study of critical events in high-stakes decisions; section VI measures the causal paths in the model compared with human-like interpretations, and conclusions and future work directions appear in section VII.

II. RELATED WORK

This section reviews the recent technologies and trends for making high-stakes decisions based on machine learning while highlighting the limitations that high-stakes decision-making must address.

A. HIGH-STAKES DECISION MAKING

High-stakes decision-making aims to prevent the worsening of a situation such as the occurrence of serious injuries and death during a first-aid incident [10]. However, the process is limited by incomplete, insufficient, and conflicting evidence, which may cause authorities to make poor decisions.

To close the gap, research on first-aid decision-making has paid attention to event descriptions using big data [11] and machine learning [12]. Devaraj [13] and Madichetty [14] used machine learning to identify requests for urgent help in critical conditions, while Sarkar *et al*. [15] predicted injury severity.

Yu *et al.* [16] employed a case-based reasoning system for supplying timed information to help authorities. Kuo *et al*[17] utilized machine learning for time predictions and identified time as a key factor in high-stakes decision-making. Yu *et al.* [18] applied machine learning to identify susceptible areas related to a natural disaster, while Zhao *et al* [19] examined locations such as public buildings in man-made disasters that affect decision making. Clearly, Spatio-temporal analysis plays a key role in high-stakes events [20], [21].

Although these studies identified factors that are helpful for high-stakes decision-making, none of them proposed meaningful relationships among those factors to aid deep explanation.

B. BAYESIAN NETWORKS

Bayesian Networks (BNs) model is an interpretable probabilistic machine learning approach. It interprets causal effect relationships using conditional dependence structure between random variables based on Directed Acyclic Graph (DAG). DAG lets agents predict the outcomes and explain how and why the results are made plausibly.

Zhou *et al.* [23] proposed the BNs model to generate if-then rules to assess risks of shipping service. Moreira *et al.* [24] proposed a BNs-based approach as an explainable model for providing insightful information in decision-making. Although these studies claimed their models provided practical explanations, they did not consider information to explain how interventions could change outcomes to support high-stake decision-makers [3]. For example, high-stakes decisions require knowing how to provide the assistance requested by an injured victim, considering where and when the event happened. These interpretations help authorities plan and respond to conditions appropriately.

Uncovering such hidden knowledge requires critical thinking, a fundamental human principle for synthesizing knowledge intelligently. It is a capability of Causal AI that software agents must imitate to model human-like intelligence. It is a challenge in Causal AI to apply that concept at the implementational level, which is still an infant in recent AI applications [25].

C. CRITICAL THINKING

Critical thinking is the requirement for supporting scientific event explanations. General critical thinking typically uses 5W1H (***Who***, ***What***, ***Where***, ***When***, ***Why***, and ***How***) to extract and describe events. For example, Yu *et al* [20] investigated event detection to support decisions. Sahoh and Choksuriwong [22] and Abebe *et al.* [23] employed a semantics-aware

event-based approach and discussed how critical thinking could be utilized in intelligent high-stakes systems. Xu *et al.* [24] proposed heuristic-based event descriptions using critical thinking for detection. However, these studies did not consider the interpretation of the circumstances that led to catastrophe. Event interpretation requires answers to ***Why*** and ***How*** questions which necessitate the use of high-level cognition to describe the events from the viewpoint of human-like intelligence.

Pearl and Mackenzie [25] defined ***How*** as interventional questions where software agents are asked to describe their reasons (e.g., *how did the critical event happen?*). ***Why*** are counterfactual questions where software agents must interpret contrastive events (e.g., *why not a different event?*). These kinds of questions are outside the bounds of our current literature although they are very important. For example, the current evidence posits that there is around a 1% chance of a catastrophic incident, but when it does occur the impact will affect a high-density population zone. Clearly, authorities should ask for reasonable deep knowledge so they can take a proactive approach to protect their citizens. Unfortunately, current critical event description approaches cannot answer ***Why*** and ***How*** questions. Instead, the burden is passed to the decision-makers as additional time-consuming and labor-intensive tasks. Critical event interpretation needs a way to model ***Why*** and ***How*** answers cognitively.

Our approach aims to contribute Causal AI based on BNs that apply critical thinking concepts to provide human-like interpretation, called Causal Bayesian Networks (CBNs). Our research challenges are 1) How to model high-stakes knowledge to provide reasonable answers based on ***Why*** and ***How***?, and 2) What are the fundamental concepts for encoding human-like interpretation to construct such an approach?

III. CAUSAL BAYESIAN NETWORKS FOR HIGH-STAKES DECISIONS

CBNs satisfy causal science that aims to produce interpretable and argumentative conclusions for high-stakes decisions based on visible evidence and prior knowledge. CBNs are a core component of agent architecture that helps agent infer plausible information [26]. Causal science consists of three main concepts: 1) questions that we need to ask software agents to reach conclusions, 2) background knowledge that software agents employ as initial grounded truth, and 3) evidence that software agents can obtain from the environment [25]. The general components of causal science are shown in Figure 1.

Figure 1 has three main elements: 1) evidence (***E***) taken from the real-world environment, 2) knowledge (***K***) encoding prior experience for plausibly interpreting the evidence, and 3) desirable conclusions (***C***) generated to answer the questions. Causal science plays a key role in connecting the real world to stakeholders because it can be applied with ***Why*** and ***How*** critical thinking to serve high-stakes decision-making. This section will explore several technologies based on causal science for modeling high-stakes decisions.

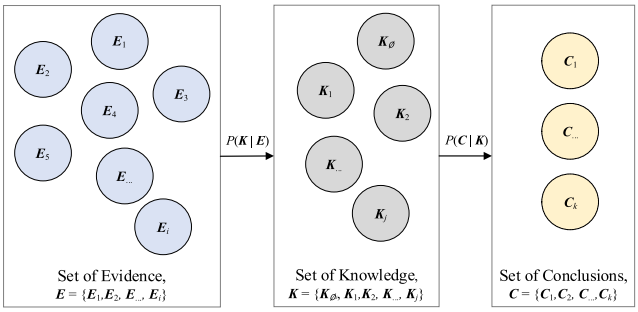


FIGURE 1. Drawing conclusions from evidence and causal knowledge.

TABLE 1. The levels of causal questions for high-stakes decisions.

Level	Plain Causal Question	Algebraic Causal Question
Association	What are the events likely to be, given the evidence? (e.g., images, texts, or videos)	Give X = x, predict Y = y $P(Y = y \mid X = x)$
Intervention	What happens if there is a shooting? And how likely is that situation?	Force X = x, predict Y = y $P(Y = y \mid do(X = x), Z)$
Counterfactual	Why was there a shooting? What would have been the effect if a shooting had not occurred based on the given evidence?	Differentiate $X = \neg x$, predict $Y_x = y_x$ $P(Y_x = y_x \mid do(X = \neg x), Z)$

A. CAUSAL QUESTIONS BASED ON CONCLUSIONS FOR HIGH-STAKES DECISIONS

Causal questions using critical thinking (5W1H) produce interpretable conclusions because good questions help people understand the chaotic real world [27]. Human thought is encoded in the form of assumptions based on human-like interpretation, as proposed by Pearl [28]. Examples of causal questions for high-stakes are shown in Table 1.

The causal questions in Table 1 are differentiated into three levels: associations, interventions, and counterfactuals. Each type is essential for software agents to mimic human-like interpretation.

Association allows a software agent to answer a related question using basic statistical conditions (e.g., detection, description, and prediction), which lets the software agent directly match the related object to exact events. For example, features such as {*gunshot*, *shooter*, *gunfire*} can match a {*shooting*} event, while features {*explosions*, *suspicious packages*, *suicide attacks*} match {*bombing*}.

Intervention is a medium-level ability that lets software agents decide on future actions. It fixes some events (e.g., $x = \{shooting\}$, $z = \{rural\ area\}$) and then interprets ***how*** future scenarios are affected (e.g., $y = \{basic\ medical\ first\ aid\}$). Intervention happens daily when authorities need to understand upcoming trends. It allows software agents to mimic human-like thinking when they have to decide the best actions with the lowest uncertainty in the real world. This cannot be

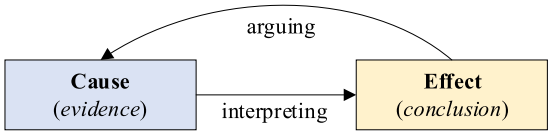


FIGURE 2. Cause-and-effect comprehension.

based on raw data alone, regardless of its size, but must also make assumptions based on cause-and-effect relationships. Software agents benefit from this by being able to simulate scenarios and present snapshots of possible futures. These are utilized by the authorities for early planning, the issuing of warnings, and preventive measures.

A counterfactual is a high-level ability that relies on human imagination. It cannot be derived from associations or interventions because the situation has not happened. For example, given the current situation $x = \{shooting\}$ we might ask "what would be (y_x) if x was *bombing*$(\neg x)$, and it happened in *a crowded area*?". This ability is needed so that software agents can adapt themselves to inexperienced situations.

Both interventions and counterfactuals go beyond traditional AI, and need human-like ability to interpret their answers. They require an explicit model based on causal relationships that can interpret both ***how*** and ***why*** answers.

B. THE CAUSAL CONCEPT FOR HIGH-STAKES DECISION

Every conclusion reached by human decision-making employs rational reasons based on knowledge and evidence [29]. Cause-and-effect comprehension is a fundamental principle for obtaining answers to causal questions that support high-stakes decisions. Cognitive comprehension is shown in the form of a simple diagram in Figure 2.

The figure shows how evidence can be used to interpret conclusions, while conclusions can argue causes. These processes are called interpretable and argumentative abilities. The argumentation provides potential causes when a conclusion has been discovered so that the software agent can answer "*What was the cause of the emergency first aids?*". The answer can be argued by finding the most possible cause of the conclusion based on prior knowledge, such as "*heavily injured victims are likely to be the cause*". On the other hand, the interpretation may provide alternative explanations of the most likely conclusion given the evidence. For example, given the evidence "*Minorly injured victims*" and the question "*What kind of first aid should be prepared?*", the software agent might answer "*It can be basic first aid because the victims are minorly injured*". The interpretation dynamically changes the answers' confidence according to new evidence.

Human beings employ cause-and-effect comprehension daily to exchange knowledge because it is a powerful tool for interpreting complex events and making decisions. Software agents need to mimic this ability to produce knowledge by answering critical questions.

C. CAUSAL MACHINE LEARNING

High-stakes decisions should use Causal ML to empower software agents to uncover knowledge from observational evidence and produce interpretable reasons. Fortunately, Bayes' theorem [30] can be used to support cognition by employing observational evidence to interpret and argue an event's reason. Bayes' theorem can be written as equation 1.

$$P(Effect|Cause) = \frac{P(Cause|Effect) \times P(Effect)}{P(Cause)} \quad (1)$$

The equation consists of four components: **Posterior** $P(Effect \mid Cause)$ computes a conclusion given evidence that is aligned with the interpretation process; **Likelihood**$P(Cause \mid Effect)$ computes the evidence given a conclusion that corresponds to the argumentation process; **Prior** $P(Effect)$ encodes the likelihood of an occurrence of a conclusion known from the past; **Evidence** $P(Cause)$ encodes the overall chance of new evidence without reference to the conclusion. They provide both interpretation and argumentation that serve the needs of the casual concept from topic B in section III.

Causal Bayesian Networks (CBNs) handle complex problems based on Bayes' theorem by encoding causality using a Directed Acyclic Graph [31]. The DAG models random variables as nodes, and semantic meaning between the nodes as edges with statistical dependency weights called conditional probabilities. A conditional probability lets a random variable conditionally control the state of another random variable according to a causal assumption computed by equation 2.

$$P(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n) = \prod_{i=1}^{n} P(\mathbf{X}_i|Pa(\mathbf{X}_i)) \quad (2)$$

We utilize Equation 2 to explain the causal concept from topic B in section III. $P(\boldsymbol{X_i})$ encodes the possible effects that interpret events for supporting decision-making. $Pa(\boldsymbol{X_i})$ encodes possible $\boldsymbol{X_i}$'s causes that clarify the ***how*** and ***why*** answers made with Equation 1. For example, given the evidence "*the heavily injured victims*" as $Pa(\boldsymbol{X_i})$, the authorities may ask "*What kind of first aids should be prepared?*". The answer can be computed by estimating the most possible conclusion $P(\boldsymbol{X_i})$ using Equation 2.In this way, CBNs represent human-like interpretation and are a powerful tool for supporting high-stakes decisions made by the software agents discussed in topic A in section III.

Suppose we need to understand a critical accident in order to provide first aid assistance in a high-stakes situation. Three variables are considered: Impact $\boldsymbol{I}$ (e.g., minor injury, heavy injury, or death), Severity $\boldsymbol{S}$ (e.g., very severe, severe, or not severe), and First Aid $\boldsymbol{F}$ (e.g., basic or emergency first aid). They can be encoded by causal assumptions using CBNs as diagrammed in Figure 3.

Figure 3 represents CBNs that can compute $\boldsymbol{F}$ based on two types of causal paths: a direct cause ($\boldsymbol{S}\rightarrow\boldsymbol{F}$) drawn as a solid line and an indirect cause ($\boldsymbol{I}\rightarrow\boldsymbol{F}$) drawn as a dashed line.

A direct cause captures the causal path determined by starting at nodes pointing towards ending nodes. $\boldsymbol{S}\rightarrow\boldsymbol{F}$ ($\boldsymbol{S}$ is

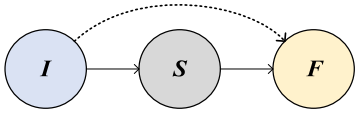


FIGURE 3. CBN-based first aid assistance for high-stakes decisions.

a cause of $\boldsymbol{F}$) which can be interpreted as accident severity, a cause that directly influences the first aid requirements. The indirect cause is a causal path determined by the unobservable evidence of intermediary nodes. In this case, only $\boldsymbol{I} = i$ is observable and influences $\boldsymbol{F}$ through $\boldsymbol{S}$, even though $\boldsymbol{S}$ is currently unobservable, and the semantic meaning for the potential effects of $\boldsymbol{F}$ is still produced. In other words, different kinds of an accident $\boldsymbol{I} = i$ may require different help $\boldsymbol{F}$, which are inferred through an intermediary variable $\boldsymbol{S}$ that is computed by equation 3.

$$\boldsymbol{F} = \int_{S} P(\boldsymbol{F}|\mathbf{S}, do(\boldsymbol{I} = i)\, P(\boldsymbol{S}|do(\boldsymbol{I} = i)\, \boldsymbol{dS} \tag{3}$$

Equation 3 expresses how a CBN uncovers possible $\boldsymbol{F}$ events when given accident impact as evidence. It shows how a software agent suffering from poor evidence (e.g., only $\boldsymbol{I} = i$) in uncertain situations is still able to compute an answer using causal understanding [32].

Causal ML can be employed for high-stakes decision making especially during critical and insecure events. This research employs Causal ML based on CBNs to encode causal assumptions and produce answers to causal questions to help make better decisions.

IV. CAUSAL ENCODING FOR HIGH-STAKES DECISIONS

The main concern of high-stakes decisions is not only to achieve the best prediction but also to understand the uncertainty factors. These express the likelihood of events that can have a devastating impact, and the interpretations of the model are a fundamental requirement for authorities when making decisions.

For example, the time ($\boldsymbol{T}$) of an accident and its location ($\boldsymbol{L}$) must be employed in first aid understanding how to access the area quickly [33]. $T = critical$ and $\boldsymbol{L} = crowded\ zone$ help the authorities interpret any difficulties by statistically associating $\boldsymbol{T}$ with $\boldsymbol{L}$ so they are likely to occur together. However, in a real-world situation, a high correlation between $\boldsymbol{T}$ and $\boldsymbol{L}$ does not mean that they directly influence one another since time does not change the location and vice versa. A hidden factor ($\boldsymbol{H}$), or confounding bias, may be invisible but can causally connect them. Software agents must learn how to model $\boldsymbol{T}$ and $\boldsymbol{L}$ bridged by $\boldsymbol{H}$ plausibly. Fortunately, causal graphs can represent this in the form of CBNs [34], [35]. The causal relationships can be encoded in several ways, and software agents can initially compute $\boldsymbol{T}$ and $\boldsymbol{L}$ dependency measured by $\boldsymbol{H}$ based on d-separation [36], as shown in Table 2.

Table 2 shows four different forms of causal graphs that encode causal assumptions when $\boldsymbol{T}$ and $\boldsymbol{L}$ are causally connected by $\boldsymbol{H}$.

The chain, inverse chain, and common cause types show that the computation of $\boldsymbol{T}$ and $\boldsymbol{L}$ is independent $\boldsymbol{T} \perp\!\!\!\perp \boldsymbol{L}$ given $\boldsymbol{H}$. In other words, if $\boldsymbol{H}$ is observed, then software agents can compute $\boldsymbol{T}$ or $\boldsymbol{L}$ depending only on $\boldsymbol{H}$ without needing further information, which is called d-separated. In contrast, if only $\boldsymbol{T}$ or $\boldsymbol{L}$ is observed, then computing them completely depends on the observation and must go through $\boldsymbol{H}$. This makes all of them dependent, which is defined as d-connected.

The collider is a special graph form whose semantic meaning contrasts with the other forms. $\boldsymbol{T}$ and $\boldsymbol{L}$ are initially independent or d-separated $\boldsymbol{T} \perp\!\!\!\perp \boldsymbol{L}$. But when $\boldsymbol{H}$ is observed, it makes $\boldsymbol{T}$ and $\boldsymbol{L}$ d-connected or $\boldsymbol{T} \not\perp\!\!\!\perp \boldsymbol{L}|\boldsymbol{H}$.

These causal graphs explain the semantic relationships between $\boldsymbol{T}$ and $\boldsymbol{L}$ through $\boldsymbol{H}$. d-separation cuts and connect the nodes in CBNs that hold the most relevant nodes for interpreting and arguing the reason for how and why it is computed. Causal graph and d-separation let software agents imitate human-like intelligence in high-stakes decision-making.

Although d-separation can encode criteria based on conditional independencies, it is unable to distinguish the semantics of how differences form between the causal graph. For example, the chain, inverse chain, and common causes types encode the same conditions (see the d-separation column in Table 2), so they cannot give significant meaning to how they semantically differ from one another.

The *do*-operation is a successor to d-separation which computes the target node by forcing some nodes to be constant. For example, the authorities may ask "*What is* $\boldsymbol{L}$ **likely to be if** $\boldsymbol{H}$*was restricted to* h?", where $\boldsymbol{L}$ is the target node and $\boldsymbol{H}$ is set to be h, which can be written as an algebraic causal question $P(\boldsymbol{L} \mid do(\boldsymbol{H} = h), \boldsymbol{T})$. The computational functions to compute the answer using different causal graph forms and a *do*-operation are shown in Table 3.

Table 3 shows the computational functions to answer $P(\boldsymbol{L} \mid do(\boldsymbol{H} = h), \boldsymbol{T})$ based on several types of causal graphs. The proof of the computational functions of $P(\boldsymbol{L}, do(\boldsymbol{H} = h), \boldsymbol{T})$ based on causal graphs is given in Appendixes. The computational functions use d-separations and*do*-operations to imitate human-like interpretations. People do not use everything that they have learned in the past to make reasonable decisions, but instead utilize the most relevant events, and neglect insignificant situations.

There are four ways to compute answers with computational functions based on graph types suitable for the problem. From the viewpoint of human-like interpretation, the ability to access an accident area and give first aid is considered. $\boldsymbol{H}$ represents the difficulty of accessing the area, and $\boldsymbol{T}$ and $\boldsymbol{L}$ represent contexts. Generally, $\boldsymbol{T}$ and $\boldsymbol{L}$ do not affect one another (e.g., period time does not influence the choice of the area) and $\boldsymbol{H}$ factors do not influence either of them. Unless the $\boldsymbol{H}$ factors are observed, $\boldsymbol{T}$ and $\boldsymbol{L}$ will plausibly influence one another (e.g., If $\boldsymbol{H} = very\ difficult$, then either $\boldsymbol{T} = critical$ or $\boldsymbol{L} = crowded\ zone$ must be true). Software agents can imitate

TABLE 2. Causal relationships between T and L connected by H.

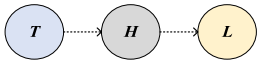
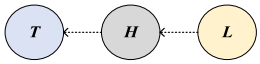
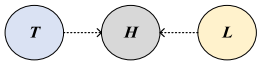
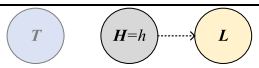
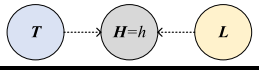
Graph Type	Triple-based-Graph	*d*-separation	Causal Assumption Description
Chain		$T \perp\!\!\!\perp L \mid H$	T and L co-occur with H so that T and L are marginally dependent through H. This means that if we know H no matter what T is, then it does not change L. This is called a causal chain.
Inverse Chain		$T \perp\!\!\!\perp L \mid H$	The graph is inversely conditional upon the causal chain.
Common Cause		$T \perp\!\!\!\perp L \mid H$	T and L co-occur with H, and H independently influences both T and L. This is called a common cause.
Collider		$T \perp\!\!\!\perp L \mid H$	T and L co-occur with H but T and L are marginally associated given H or its descendants, otherwise, they are completely independent. This is called a collider.

TABLE 3. Computational Functions of L based on a do-operation by fixing H = h.

Graph Type	Triple-based-Graph by fixing $H = h$	Computational Function
Chain		$P(L \mid do(H = h), T) = P(L \mid do(H = h))$
Inverse Chain		$P(L \mid do(H = h), T) = P(L)$
Common Cause		$P(L \mid do(H = h), T) = P(L \mid do(H = h))$
Collider		$P(L \mid do(H = h), T) = P(L)$

such understanding by considering the semantic relationships between T, L, and H through d-separation and do-operation. In particular, H is fixed to determine the (in)dependence between T and L. Also, the function that computes the answer is unlikely to be a chain, inverse chain, or common cause, because the collider-based-causal graph is more suitable.

This section has shown how software agents can imitate human-like intelligence using CBNs. CBNs semantically encode knowledge to deal with high-stakes problems which allow the agents to produce interpretable and argumentative knowledge. This research employs CBNs to construct a causal model to benefit high-stakes decision-making.

V. CASE STUDY: CAUSAL MACHINE LEARNING MODEL FOR HIGH-STAKES DECISION

The goal of this section is to develop a causal model to support decision-making in high-stakes management strategies. Therefore, the research questions are 1) what kinds of critical factors are relevant to high-stakes decision-making, 2) how to represent these factors to generate knowledge, and 3) how to approve these critical factors to causally explain events in real-world environments.

Oroszi [37] identified terrorism as a high-stakes situation, with intensive time pressure and high uncertainty, which must be handled by interpretable knowledge. Terrorism affects the well-being of people, breaks society's function, and is feared by counties around the world. Thailand is one of the top ten countries suffering from its impact [38]. From January 2004 to June 2019, Thailand had to deal with 20,323 terrorist attacks, with 6,997 people killed and 13,143 injured, as reported by Deep South Watch [39]. As a result, we have chosen Thailand's terrorism as an environment in which to build a high-stakes decision-making model.

TABLE 4. The critical factors considered in high-stakes issues.

Critical Factor	Factor's State	Critical Thinking based-Description
Time (***T***)	*critical*, *difficult*, *normal*	When an event happens (***When***).
Location (***L***)	*very-crowded*, *crowded*, *not crowded*	The specific location of an event (***Where***).
Accident (***A***)	*bombing*, *shooting*, *battling*	An event type (***What***).
Impact (***I***)	*mortality, injury*, *safe*	A measure of the victim's lifesaving level (***Who***).
Search and Rescue (***SR***)	*crucial search*, *difficult search*, *normal search*	The capability of the rescue personnel to deal with the event (***How***).
Severity (***S***)	*very severe*, *severe*, *not severe*	The level of violence suffered by victims (***How***).
First Aid (***F***)	*immediate response*, *prepared response*, *monitoring*	How medical staff should respond to the event (***How***).

Time (T) and location (L) are general factors (as discussed in topic A in section II) for explaining the causal encoding of section IV. However, high-stakes issues require more than just time and location data to determine the trade-offs between low chance and serious consequences. Wang *et al.* [40] and Mujalli *et al.* [41] argued that accident types and their impact are also important factors for decision-making. Based on the literature, we have highlighted the following critical factors in Table 4 to be represented by random variables.

Table 4 shows critical factors considered as random variables aligned with human critical thinking (5W1H). The variables can be categorized into dependent and independent groups. The **First Aid** dependent variable will be changed during the experiment depending on other factors. The other variables are independent whose states can randomly occur and control the dependent variable. In the other words, **First Aid**'s states are exposed when the independent variable's

states are observed or measured. For example, the experiment can set **First Aid** = *immediate response* if we observe **Accident** = *bombing*, to compute the odds of *the immediate response* given the *bombing* that has occurred. This is similar to how people interpret a situation on an everyday basis. However, all the values of the random variables are determined based on qualitative and abstract understanding. This digitization of human-like intelligence to support high-stakes decision-making scenarios is a challenge.

The digitization of human-like intelligence focuses on measuring the corresponding between the independent variables ($\boldsymbol{X}$) and the dependent variable ($\boldsymbol{Y}$). Conditional probability is employed to observe how likely that the $\boldsymbol{Y}$ states occur given states of $\boldsymbol{X}$. This can be symbolized by $P(\boldsymbol{Y} \mid \boldsymbol{X})$ where $\boldsymbol{Y} = \{$**First Aid**$\}$ and $\boldsymbol{X} = \{$**Time**, **Location**, **Accident**, **Impact**, **Search and Rescue**, **Severity**$\}$. The basic hypothesis is that if event $\boldsymbol{X} = x$ and event $\boldsymbol{Y} = y$ are mutually relevant, then the conditional probability between them can be represented by a Conditional Probability Table (CPT). The matrix for $P(\boldsymbol{Y} = y_i | \boldsymbol{X} = x_j)$ can be formed by CPT_{ij} where i is the range of the independent variable and j is the range of the dependent variable. This matrix based on CPT_{ij} must compute the column values using the $\sum_j \text{CPT}_{ij} \approx 1$ for all i.

We employed 20,323 terrorism events taken from the Deep South Watch Database [42] to perform the CPT of $P(\boldsymbol{Y} \mid \boldsymbol{X})$. The probability outcomes of the dependent variable under the condition of independent variables were represented as a matrix-based CPT with colored graphics using a heatmap visualization to show the likelihood of events co-occurring (red to white). The associations between the $\boldsymbol{X}$ and $\boldsymbol{Y}$ sets are shown in Table 5.

Table 5 displays how likely event $\boldsymbol{Y}$ is given event $\boldsymbol{X}$ using a color scale that ranges from red for higher probabilities to white for lower probabilities. For example, the conditional probability of $P($**First Aid** = *immediate* | **Location** = *very-crowded*$)$ is 72% and $P($**First Aid** = *monitoring* | **Location** = *not crowded*$)$ is 80%, which means they are highly likely to co-occur. Clearly, when a "*very-crowded*" area is observed, an "*immediate*" response should be considered by authorities.

Although the pattern-based CPT represents the probabilities of two events co-occurring, it is a correlated relationship and does not plausibly signify causation between $\boldsymbol{X}$ and $\boldsymbol{Y}$. For instance, in the above example, human intuition can understand that certain very-crowded areas, such as parks, casinos, and shopping malls, do not require an immediate response from first aid services (e.g., blood reserves, breathing apparatus, and recovery vehicles). This shows that relevant decisions depend on hidden factors that need to be encoded. In other words, causal science is required to model high-stakes events that require interpretations and arguments rather than purely highly correlated scores.

A. CAUSAL EFFECT MODELLING

The goal of this section is to determine causal relationships between random variables from Table 4 by imitating human commonsense to encode transparent and testable knowledge. For example, we discussed the semantic relationships between **Location** ($\boldsymbol{L}$), **Time** ($\boldsymbol{T}$), and **Search and Rescue** ($\boldsymbol{SR}$) in section IV as represented by a collider-based causal graph that can be understood by software agents.

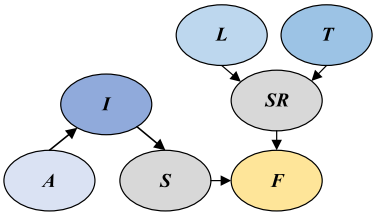


FIGURE 4. Causal graph CBNs encoded Human Understanding for high-stakes decisions.

Accident ($\boldsymbol{A}$), **Impact** ($\boldsymbol{I}$), and**Severity** ($\boldsymbol{S}$) do not randomly occur, since $\boldsymbol{A}$ might directly cause $\boldsymbol{I}$ while $\boldsymbol{I}$ influences $\boldsymbol{S}$ so that $\boldsymbol{A}$ indirectly convinces $\boldsymbol{S}$. We can represent these causal relationships using a causal chain.

$\boldsymbol{SR}$ and $\boldsymbol{S}$ are usually independent except when the authorities ask for the likelihood of **First Aid** ($\boldsymbol{F}$), which will cause $\boldsymbol{SR}$ and $\boldsymbol{S}$ to influence one another. Relationships of these types can be discovered by observation (e.g., via statistical studies or by talking to experts). We draw these causal assumptions as the CBNs shown in Figure 4.

The model represents how $\boldsymbol{L}$, $\boldsymbol{T}$, and $\boldsymbol{A}$ are the root causes in the graph and are d-separated, which occur independently in physical reality because they connect in the form of a collider-based causal graph. This means that the accident can happen anywhere and at any time unless the authorities ask about $\boldsymbol{F}$, and all of them become d-connected. For example, consider when the software agent observes that time $\boldsymbol{T}$ = *normal* but the situation response is $\boldsymbol{F}$ = *immediate*. The causal model generates the rational reason for this situation that $\boldsymbol{S}$ is most likely to be *very severe*, probably due to $\boldsymbol{I}$ = *mortality*, and the most possible reason is $\boldsymbol{A}$ = *bombing*. This is because its impact can affect a high-density population zone that sets $\boldsymbol{SR}$ = *difficult search* since $\boldsymbol{L}$ = *crowded* area is highly likely. This shows how $\boldsymbol{L}$, $\boldsymbol{T}$, and $\boldsymbol{A}$ are linked when $\boldsymbol{F}$ is aligned with a d-connected graph.

An additional aspect is that $\boldsymbol{SR}$ and $\boldsymbol{S}$ directly cause $\boldsymbol{F}$, which is independent of the other factors. For example, given $\boldsymbol{I}$ = *mortality* the authorities may ask how likely $\boldsymbol{F}$ will be. The graph shows that $\boldsymbol{I}$ and $\boldsymbol{F}$ do not directly influence each other but provides an intuition that $\boldsymbol{I}$ = *mortality* may cause $\boldsymbol{S}$ = *very severe* which must trigger $\boldsymbol{F}$ = *immediate*. However, if $\boldsymbol{S}$ = *very severe* is observed, then it directly influences $\boldsymbol{F}$ = *immediate* without requiring any information from $\boldsymbol{I}$ because $\boldsymbol{S}$ already summarizes $\boldsymbol{I}$. This allows the software agent to interpret it gives a reasonable conclusion from the viewpoint of the causal model.

A causal model can interpret conclusions from high-stake situations for two main reasons: 1) all the causal relationships can be interpreted by a software agent and generate ***How*** and ***Why*** answers for decision-makers, and 2) relationships in the

TABLE 5. The CPT of dependent variables conditioned upon independent variables.

Independent / Dependent	Location			Time			Search			Impact			Severity			Accident			
First Aid	*very-crowded*	*crowded*	*not-crowded*	*critical*	*difficult*	*normal*	*critical*	*difficult*	*normal*	*mortality*	*injury*	*safe*	*very severe*	*severe*	*not severe*	*battle*	*bombing*	*shooting*	*turbulence*
immediate	0.72	0.11	0.17	0.14	0.67	0.19	0.67	0.33	0.00	1.00	0.00	0.00	0.75	0.23	0.02	0.17	0.78	0.05	0.00
prepared	0.26	0.47	0.28	0.28	0.58	0.14	0.39	0.43	0.19	0.43	0.34	0.23	0.48	0.27	0.25	0.08	0.67	0.23	0.02
monitoring	0.01	0.19	0.80	0.08	0.55	0.37	0.00	0.24	0.75	0.00	0.11	0.89	0.08	0.40	.52	0.11	0.35	0.38	0.16

causal model are transparent which can be troubleshot by the authorities and experts to address how the model wrongly connects variables. If the answer from the model is conflicted, the model's structure can be easily revised and updated by experts.

The causal model can encode commonsense to support high-stake decision-making based on qualitative design. This requires the real-world environment to be specified so that observational data can be collected to fit the parameters and transform the models into a quantitative representation that can be evaluated using model fitting.

B. DATA PREPROCESSING

Twitter [12] allows software agents to consume real-time and worldwide observations, and so can be utilized to collect information on dependent and independent variables. However, most tweets consist of unimportant words, symbols, conjunctions, and abbreviations, so natural language technology is needed to handle such problems. Our information extraction technique is aligned with similar approaches [43], [44] that detect states of random variables from tweets. The overview architecture is shown in Figure 5.

Figure 5 shows that there are five main components for the extraction of random variables and their states from Twitter. In 1), Tweet streaming collects real-time tweets as text and feeds them into 2) Tokenization and Noise Removal to split the strings into tokens and remove insignificant values. In stage 3), Named Entity Recognition identifies the meanings of tokens using a Human Critical Thinking Model [22]. This identifies elements such as people, building and place, time, and accident. The states are then matched with variables by employing the resulting contexts in 4) Variable and State Description. Statements are generated from this information by in stage 5) Variable and State Creation.

To show the difference between the input (raw tweet) and output (information) of the information extraction process, we sampled a tweet posted by @TichilaThaipbs, a field reporter for Thai PBS (Thai Public Broadcasting Service). Her tweet was evaluated by two emergency medicine physicians and three practitioners from Prince of Songkla University Hospita; they confirmed that the tweet showed that the situation called for immediate first aid. In addition, the tweet was converted into a form suitable for the software agent by our extraction process, resulting in Table 6.

TABLE 6. The variables and state information extracted from a tweet using our approach.

Raw Tweet (Input)	"*Bombing around the highway road in Pa-bang village, Thepa District, Songkla Province. One soldier has been killed and six injured. Reported at 12:38 PM, Oct 1, 2020*".
Information (Output)	[Bombing <*bomb* ∈ **Accident**> around highway road < *highway road* ∈ **Location**> in Pak-Bang village <*village* ∈ **Location**>, Thea-pa district <*district* ∈ **Location**>, Songkhla province <*crowded* ∈ **Location** >. One solder < *victim* ∈ **Impact**> has been killed and six injured < *one death, six injured* ∈ **Impact**> reported at 12:38 PM Oct 1, 2020 <*12:38 PM (critical) Oct 1, 2020* (*Thai holiday*) = *crucial* ∈ **Time**>]

The second row of Table 6 gives the tweet with semantic tags with extracted information in a software agent readable format. Light-gray text marks the tokens considered to be noise, while the red text is words that may denote states and variables. The italic text is states, and the bold text is variables. These are linked by subset (∈) to denote the context of the texts in the form of state and variable information which will become observational evidence used by the causal model. In the next section, we propose a methodology for measuring the causal model with this evidence.

VI. EXPERIMENT SETUP

One of the best-known abilities of the causal machine learning model is its predictive ability. However, human-like intelligence extends far beyond pure prediction. Its most challenging aspect is how to evaluate the plausibility of knowledge in causal paths because testing this is a difficult task with no standard tool to measure its performance and needs. This is key for allowing the model to interpret and argue about the reasons for high-stake decision-making.

The motivation of our experiment is to measure the rationality of causal paths in a DAG using observed data as evidence. Intuitions are proved by interpreting causality with d-separations that state whether the relationships between variables are separated or connected.

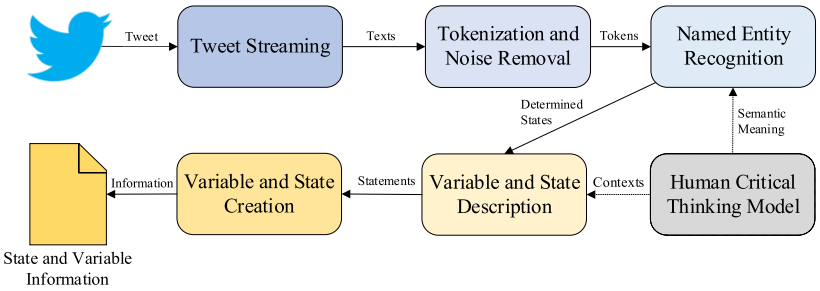


FIGURE 5. The architecture of information extraction approach.

A. MEASUREMENT METRICS

Causal Odds Ratio (Causal OR) measures the robust causality between random variables. It is written as Causal OR($\boldsymbol{X}$, $\boldsymbol{Y}$ | $\boldsymbol{Z}$), where $\boldsymbol{Y}$ is a set of dependent variables, $\boldsymbol{X}$ is a set of independent variables, and $\boldsymbol{Z}$ is a set of confounding bias variables. It measures how event $\boldsymbol{X} = x$ can influence event $\boldsymbol{Y} = y$ conditioned on event $\boldsymbol{Z} = z$. Causal OR has been proposed as the basis for the *do*-operation [45], and we use it to measure our assumptions about the causal graph:

$\boldsymbol{X} = x$ is an event of interest that must be measured for robust causality with $\boldsymbol{Y} = y$ while $\boldsymbol{X} = \neg x$ are the rest of the events that oppose to $\boldsymbol{X} = x$. $\boldsymbol{Y} = \neg y$ is a reference set by the majority of the sample events that is employed to calculate the ratio between $\boldsymbol{X} = x$ and $\boldsymbol{Y} = y$ normalized by $\boldsymbol{Y} = \neg y$. The Causal OR results can be semantically interpreted as:

- Causal OR ≈ 1: $\boldsymbol{X} = x$ does not change the likelihood of $\boldsymbol{Y} = y$ given $\boldsymbol{Z} = z$, (i.e., $\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y}|\boldsymbol{Z}$)
- Causal OR > 1: $\boldsymbol{X} = x$ increases the likelihood of $\boldsymbol{Y} = y$ given $\boldsymbol{Z} = z$, (i.e., $\boldsymbol{X} \not\perp\!\!\!\perp \boldsymbol{Y}|\boldsymbol{Z}$)
- Causal OR < 1: $\boldsymbol{X} = x$ decreases the likelihood of $\boldsymbol{Y} = y$ given $\boldsymbol{Z} = z$, (i.e., $\boldsymbol{X} \not\perp\!\!\!\perp \boldsymbol{Y}|\boldsymbol{Z}$).

Causal OR shows the strength of the relationships between variables but cannot confirm whether they co-occur by chance or have statistical significance.

The Causal P-value is a probability score to express a significance under the causal assumption of $\boldsymbol{X}$ and $\boldsymbol{Y}$ conditional on $\boldsymbol{Z}$. A Causal P-value of less than 0.001 is considered of high significance which means that all evidence lower than 1 out of 1000 that $\boldsymbol{X}$ and $\boldsymbol{Y}$ co-occurred by random chance given $\boldsymbol{Z}$.

The Causal Confident Interval (Causal CI) measures the precision of Causal OR of $\boldsymbol{X}$ and $\boldsymbol{Y}$ conditional on $\boldsymbol{Z}$. A Causal CI of 95% states that the range of the Causal OR of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is sure that its relevant evidence lies within 95% of all evidence.

B. DATASET

We collected evidence from field reporters to determine relevant events for all the random variables using the Twitter platform between May 16 and June 16, 2019, resulting in 100,000 transactions. These collected data were labeled by experts from Deep South Watch, the center of conflict studies based on the national security-related decision. The resulting variables and states were employed to estimate hyper-parameters in the proposed CBNs model structure that we mentioned in section V using an Expectation-Maximization algorithm [46]. The CBNs model were utilized to measure event interpretation's cause-and-effect relationships.

C. RESULTS

We measured the causal model as detailed in topic B from section V to examine the relationships between two variables (e.g., between $\boldsymbol{X}$ and $\boldsymbol{Y}$) given a third variable (e.g., a certain value of $\boldsymbol{Z}$) to determine them to be d-separated (e.g., $\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}$) or d-connected (e.g., $\boldsymbol{X} \not\perp\!\!\!\perp \boldsymbol{Y} \mid \boldsymbol{Z}$). We broke the model into sub-graphs in the form of a triple-based graph to measure its semantic relationship. The first triple-based graph uses **Location** ($\boldsymbol{L}$), **Time** ($\boldsymbol{T}$), and **Search and Rescue** ($\boldsymbol{SR}$), the second triple-based graph employs **Accident** ($\boldsymbol{A}$), **Severity** ($\boldsymbol{S}$), and **Impact** ($\boldsymbol{I}$), the third utilizes **First Aid** ($\boldsymbol{F}$), $\boldsymbol{S}$, and $\boldsymbol{SR}$, and the last uses $\boldsymbol{T}$, $\boldsymbol{L}$, and $\boldsymbol{F}$.

1) SUB-GRAPH 1

Given $\boldsymbol{L} \rightarrow \boldsymbol{SR} \leftarrow \boldsymbol{T}$ or P($\boldsymbol{L}$, $\boldsymbol{T}$ | do($\boldsymbol{SR}$)), we set $\boldsymbol{L}$ as a dependent variable, $\boldsymbol{T}$ as an independent variable, and $\boldsymbol{SR}$ as a confounding bias variable. The hypothesis was "Is $\boldsymbol{T}$ necessary to interpret $\boldsymbol{L}$ given a certain state of $\boldsymbol{SR}$?"

The Expectation: the collider-based graph set $\boldsymbol{SR}$ to $do(\boldsymbol{SR} = sr)$, so $\boldsymbol{T}$ and $\boldsymbol{L}$ must be d-connected —$\boldsymbol{T} \not\perp\!\!\!\perp \boldsymbol{L} \mid do(\boldsymbol{SR} = sr)$. In contrast, if $\boldsymbol{SR}$ is unknown and sets $\boldsymbol{SR}$ to $do(\boldsymbol{SR} = marginal)$, $\boldsymbol{T}$ and $\boldsymbol{L}$ must be d-separated—$\boldsymbol{L} \perp\!\!\!\perp \boldsymbol{T} \mid do(\boldsymbol{SR} = marginal)$. Therefore, the Causal OR of the given sub-graph should converge to "1".

We measured the first sub-graph utilizing our dataset, and the Causal OR of $\boldsymbol{T}$ and $\boldsymbol{L}$ given $\boldsymbol{SR}$ is shown in Table 7.

On the right of Table 7, the Causal OR of $do(\boldsymbol{SR} = marginal)$ is approximately 1, which means that $\boldsymbol{L}$ and $\boldsymbol{T}$ are generally d-separated without considering $\boldsymbol{SR}$. The Causal P-Value verifies that the evidence of $\boldsymbol{T} = difficult$ is around 68% by random chance which means that the relationship between $\boldsymbol{L}$ and $\boldsymbol{T}$ is not statistically significant. Moreover, the Causal OR is in the Causal CI range of both $\boldsymbol{T} = difficult$ (i.e., 0.983-1.027) and $\boldsymbol{T} = critical$ (i.e., 0.967-1.005). These measured precisions for the d-separated relationship are consistent with our expectations.

TABLE 7. Measuring L → SR ← T using Causal OR by setting T = *normal* to the reference event.

Interpreted L	SR = *crucial*		SR = *difficult*		SR = *normal*		SR = *marginal*	
	T = *critical*	T = *difficult*	T = *critical*	T = *difficult*	T = *critical*	T = *difficult*	T = *critical*	T = *difficult*
Causal OR	1.416	2.593	0.147	0.307	0.916	0.535	**0.986**	**1.004**
Causal P-Value	<0.001	<0.001	<0.001	<0.001	0.001	<0.001	**0.677**	**0.162**
Causal CI-95%	1.290-1.554	2.358-2.852	0.138-0.156	0.291-0.323	0.868-0.966	0.519-0.551	**0.967-1.005**	**0.983-1.027**

TABLE 8. Measuring L → F ← T using Causal OR by setting T = *normal* to the reference event.

Interpreted L	F = *immediate response*		F = *prepared response*		F = *monitoring*		F = *marginal*	
	T = *critical*	T = *difficult*	T = *critical*	T = *difficult*	T = *critical*	T = *difficult*	T = *critical*	T = *difficult*
Causal OR	0.953	1.384	0.873	0.946	0.835	0.723	0.986	1.004
Causal P-Value	0.272	<0.001	<0.001	<0.001	<0.001	<0.001	0.677	0.162
Causal CI-95%	0.874-1.038	1.270-1.508	0.846-0.901	0.921-0.973	0.799-0.872	0.701-0.745	0.967-1.005	0.983-1.027

TABLE 9. Measuring A → I → S using Causal OR by setting S = *normal* to the reference event.

Interpreted A	I = *crucial*		I = *difficult*		I = *normal*		I = *marginal*	
	S = *critical*	S = *difficult*	S = *critical*	S = *difficult*	S = *critical*	S = *difficult*	S = *critical*	S = *difficult*
Causal OR	0.971	0.975	0.997	1.004	0.998	1.006	1.384	1.214
Causal P-Value	0.091	0.098	0.031	0.026	0.797	0.918	<0.001	<0.001
Causal CI-95%	0.913-1.043	0.907-1.039	0.969-1.025	0.974-1.035	0.958-1.057	0.970-1.027	1.359-1.409	1.193-1.236

TABLE 10. Measuring SR → F → S using Causal OR by setting S = *normal* to the reference event.

Interpreted SR	F = *immediate response*		F = *prepared response*		F = *monitoring*		F = *marginal*	
	S = *critical*	S = *difficult*	S = *critical*	S = *difficult*	S = *critical*	S = *difficult*	S = *critical*	S = *difficult*
Causal OR	1.785	1.161	0.334	0.636	0.909	0.740	1.007	1.011
Causal P-Value	<0.001	0.011	<0.001	<0.001	0.001	<0.001	0.414	0.246
Causal CI-95%	1.606-1.983	1.034-1.302	0.323-0.345	0.616-0.657	0.860-0.960	0.716-0.765	0.989- 1.027	0.992-1.031

In the case of a given state of SR, the rest of the table shows that the Causal OR of $do(SR = crucial, difficult, normal)$ is far from 1, which induces both L and T to be d-connected. This is verified by the Causal P-Value being less than 1% by random chance and 1 is not in the Causal CI ranges.

In summary, our causal encoding of L, T, and SR using a collider-based graph fits the real-world evidence.

2) SUB-GRAPH 2

Given $L \rightarrow F \leftarrow T$ where F is a child of SR that is also considered a collider-based graph, we set L as a dependent variable, T as an independent variable, and F as a confounding bias variable. The hypothesis was "Is T necessary to interpret L given a certain state of F?"

The Expectation: the collider-based graph set F to $do(F = f)$, so T and L must be d-connected—$T \not\perp\!\!\!\perp L \mid do(F = f)$. This suggests the same trend as the first hypothesis of sub-graph 1.

We measured this hypothesis utilizing our dataset, and the Causal OR of T and L given F is shown in Table 8.

Table 8 shows that when F is set to a constant, the Causal OR of L and T displays the same trends as Table 7. However, in the case of the Causal OR of $T = critical$, it is slightly different because of increased uncertainty. For example, the Causal P-Value of $T = critical$ given $F = immediate\ response$ is around 27% by random chance while "1" is in the Causal CI range (i.e., 0.874-1.038) because F is indirect evidence of L and T and connects through SR. It reduces confidence as human understanding in the same way when we observe indirect evidence that confirms our belief less well than direct observations.

3) SUB-GRAPH 3

Given $A \rightarrow I \rightarrow S$, we set A as a dependent variable, S as an independent variable, and I as a confounding bias variable. The hypothesis was "Is S necessary to interpret A given a certain state of I?"

$$\text{Causal OR}(X, Y|Z) = \frac{P(Y = y|do(X = x), Z = z)P(Y = \neg y|do(X = \neg x), Z = z)}{P(Y = y|do(X = \neg x), Z = z)P(Y = \neg y|do(X = x), Z = z)} \tag{4}$$

The Expectation: the chain-based graph set $\boldsymbol{I}$ to $do(\boldsymbol{I} = i)$, so $\boldsymbol{A}$ and $\boldsymbol{S}$ must be d-separated—$\boldsymbol{A} \not\perp\!\!\!\perp \mathbf{S} \mid do(\boldsymbol{I} = i)$. Therefore, the Causal OR of the chain-based graph given $do(\boldsymbol{I} = i)$ should converge to 1. In contrast, if $\boldsymbol{I}$ is undetermined by setting $\boldsymbol{I}$ to $do(\boldsymbol{I} = marginal)$, the relationship must be d-connected between $\boldsymbol{A}$ and $\boldsymbol{S}$—$\boldsymbol{A} \not\perp\!\!\!\perp \boldsymbol{S} \mid do(\boldsymbol{I} = marginal)$.

We measured this hypothesis utilizing our dataset, and the Causal OR of $\boldsymbol{A}$ and $\boldsymbol{S}$ given $\boldsymbol{I}$ is shown in Table 9.

On the right of Table 9, the Causal OR of $\boldsymbol{A}$ and $\boldsymbol{S}$ given $d0o(\boldsymbol{I} = marginal)$ is for from 1, which means that they are d-connected. This is verified by 1 not being in the Causal CI range (i.e., 1.359-1.409 for $\boldsymbol{S} = critical$ and 1.193-1.236 for $\boldsymbol{S} = difficult$) and the Causal P-Value shows that less than 1% of the evidence can occur by random chance. In other words, if the evidence from $\boldsymbol{I}$ is unobserved, the knowledge of $\boldsymbol{S}$ must be summarized from $\boldsymbol{A}$ as indirect evidence. While the rest of the table shows that the Causal OR of $\boldsymbol{A}$ and $\boldsymbol{S}$ given $do(\boldsymbol{I} = crucial, difficult,$ and $normal)$ is approximately 1, which means that they are d-separated. This is similar to human cognitive understanding when the knowledge that $\boldsymbol{I}$ summarizes $\boldsymbol{A}$ means that $\boldsymbol{A}$ is not an important factor for interpreting $\boldsymbol{S}$.

4) SUB-GRAPH 4

Given $\boldsymbol{SR} \rightarrow \boldsymbol{F} \leftarrow \boldsymbol{S}$, we set $\boldsymbol{SR}$ as a dependent variable, $\boldsymbol{S}$ as an independent variable, and $\boldsymbol{F}$ confounding bias variable. The hypothesis was "Is $\boldsymbol{S}$ necessary to interpret $\boldsymbol{SR}$ given a certain state of $\boldsymbol{F}$?"

The Expectation: the collider-based graph set $\boldsymbol{F}$ to $do(\boldsymbol{F} = f)$, so $\boldsymbol{SR}$ and $\boldsymbol{S}$ must be d-connected—$\boldsymbol{SR} \not\perp\!\!\!\perp \boldsymbol{S} \mid do(\boldsymbol{F} = f)$. The expectation will display the same trend as the hypotheses in sub-graphs 1 and 2.

We measured this hypothesis utilizing our dataset, and the Causal OR of $\boldsymbol{SR}$ and $\boldsymbol{S}$ given $\boldsymbol{F}$ is shown in Table 10.

On the right of Table 10 $\boldsymbol{SR} \perp\!\!\!\perp \boldsymbol{S} \mid do(\boldsymbol{F} = marginal)$ is approximately 1, which means $\boldsymbol{SR}$ and $\boldsymbol{S}$ are d-separated if $\boldsymbol{F}$ is unexplored. The rest depends on each other semantically when $\boldsymbol{F}$ is given.

D. DISCUSSIONS

The proposed CBNs in Section VI-C let software agents break and choose the relevant variables to infer knowledge based on d-connected and d-separated. The structure is cause-and-effect relationships, dynamically based on evidence to determine the related random variables, cut off unrelated, and produce high-stakes knowledge. For example, the question is, "What is a probability of search and rescue ($\mathbf{SR}$) can be trouble ($\mathbf{SR} = difficult$) given the incident period is in the morning?". CBNs employ the incident period as critical time ($\mathbf{T} = critical$) according to the sense about rush hour. The location ($\mathbf{L}$) is an unobserved variable. However, agents can still inference $\mathbf{L}$ based on marginal distribution because the high-stakes knowledge from subgraph 1 shows that $\mathbf{SR}$ information causally depends upon $\mathbf{T}$ and $\mathbf{L}$. In contrast, software agents do not accumulate $\mathbf{A}$, $\mathbf{I}$, and $\mathbf{S}$ in the inference process because CBNs let them understand which variables are useless or useful based on d-connected and d-separated in high-stakes situations.

CBNs help software agents deal with insufficient evidence because they can compute both direct and indirect evidence. Indirect evidence is often considered outliers because of uncertainty and therefore excluded from the model. Although indirect evidence may produce an unclear outcome, it is still helpful if software agents can explain how and why such effects are made. This shows that the CBNs can encode human-like sophisticated knowledge, especially in sensitive cases of high-stakes events.

Moreover, tables 8-10 show that the same dataset provides different facts when setting the confounding bias variables to be constant. The problem is known as Simpson's paradox [31], and only experts could explain how and why it happens. The paradox may confuse non-expert decision-makers and cause difficulty in the high-stakes decision-making process. The CBNs help software agents realize Simpson's paradox and deal with high-stakes situations effectively.

VII. CONCLUSION

High-stakes decision-making deals with highly uncertain events that have a low chance (of occurring) but have a high impact when they do. Interpretable knowledge is required to understand events to prevent bad outcomes.

This research used Causal AI for high-stakes decision-making by utilizing causal science to encode human-like intelligence. Causal encoding based on d-separation and do-operation was applied to model causal assumptions as represented by CBNs with Causal OR, Causal P-Value, and Causal CI used to discover causal effects by measuring the commonsense behind a graph. Causal OR measured the robustness of the causality between random variables, Causal P-Value measured if the Causal OR occurred with statistical significance, and Causal CI confirmed whether the Causal OR was precisely aligned with the evidence. Our experiment shows that CBNs can encode commonsense based on causal assumptions by measuring their rationality using observed data as evidence. The results confirm that employing a causal model can add a significant level of cognitive understanding to high-stakes decision-making.

In the future, we plan to develop an automatic mechanism to generate causal assumptions based on unknown scenarios. This is needed when the model is applied to a new environment and needs to evolve according to new evidence. We hope to enhance the model's flexibility by employing variational inference to generate potential samples for estimating causal paths. This will allow the model to learn unknown events in high-stakes situations.

APPENDIXES

These four Appendixes give the functions based on d-separation and do-operation for whether two variables are semantically dependent or independent given a confounding bias variable. They employ a conditional (in)dependent concept based on the chain rule where $P(\boldsymbol{X}_1, \boldsymbol{X}_2 \ldots \boldsymbol{X}_i) =$

$P(\boldsymbol{X}_i|\boldsymbol{X}_1 \ldots \boldsymbol{X}_{i-1})$ transforms into $P(\boldsymbol{X}_1, \boldsymbol{X}_2 \ldots \boldsymbol{X}_i) = P(\boldsymbol{X}_i|Pa(\boldsymbol{X}_i))$ using CBNs.

According to Table 3 in section IV, there are four types of causal graphs: causal chain, inverse causal chain, common cause, and collider. Each of them consists of three variables: $\boldsymbol{T}, \boldsymbol{H}$, and $\boldsymbol{L}$, with $\boldsymbol{T}$ the dependent variable, $\boldsymbol{L}$ an independent variable, and $\boldsymbol{H}$ the confounding bias variable. The hypothesis is "Is $\boldsymbol{T}$ necessary to compute $\boldsymbol{L}$ given a certain state of $\boldsymbol{H}$?".

APPENDIX A

Given the causal chain, $\boldsymbol{T} \rightarrow \boldsymbol{H} \rightarrow \boldsymbol{L}$, the $\boldsymbol{L}$ function can be expressed as:

$$\begin{aligned} &P\left(\boldsymbol{L} \mid do(\boldsymbol{H}=h), \boldsymbol{T}\right) \\ &= \frac{P(\boldsymbol{L}|do(\boldsymbol{H}=h))P(do(\boldsymbol{H}=h)|\boldsymbol{T})P(\boldsymbol{T})}{P(do(\boldsymbol{H}=h), \boldsymbol{T})P(\boldsymbol{T})} \\ &= P(\boldsymbol{L}|do(\boldsymbol{H}=h)) \end{aligned} \tag{5}$$

The outcome is $P(\boldsymbol{L}|do(\boldsymbol{H}=h), \boldsymbol{T}) = P(\boldsymbol{L}|do(\boldsymbol{H}=h))$, which means that the $\boldsymbol{L}$ function is d-separated from $\boldsymbol{T}$ when $\boldsymbol{H}$ is set to $do(\mathbf{H}=h)$.

APPENDIX B

Given the inverse causal chain, $\boldsymbol{T} \leftarrow \boldsymbol{H} \leftarrow \boldsymbol{L}$, the $\boldsymbol{L}$ function can be expressed as:

$$\begin{aligned} &P\left(\boldsymbol{L} \mid do(\boldsymbol{H}=h), \boldsymbol{T}\right) \\ &= \frac{P(do(\boldsymbol{H}=h))|L)P(\boldsymbol{T}|do(\boldsymbol{H}=h))P(L)}{P(do(\boldsymbol{H}=h)|L)P(\boldsymbol{T}|do(\boldsymbol{H}=h))} \\ &= P(\boldsymbol{L}) \end{aligned} \tag{6}$$

The outcome is $P(\boldsymbol{L}|do(\boldsymbol{H}=h), \boldsymbol{T}) = P(\boldsymbol{L})$, which means that the $\boldsymbol{L}$ function is independent and d-separated from the rest.

APPENDIX C

Given the common cause, $\boldsymbol{T} \leftarrow \boldsymbol{H} \rightarrow \boldsymbol{L}$, the $\boldsymbol{L}$ function can be expressed as follows:

$$\begin{aligned} &P\left(\boldsymbol{L} \mid do(\boldsymbol{H}=h), \boldsymbol{T}\right) \\ &= \frac{P(\boldsymbol{L}|do(\boldsymbol{H}=h))P(\boldsymbol{T}|do(\boldsymbol{H}=h)P(do(\boldsymbol{H}=h))}{P(\boldsymbol{T}|do(\boldsymbol{H}=h))P(do(\boldsymbol{H}=h))} \\ &= P(\boldsymbol{L}|do(\boldsymbol{H}=h)) \end{aligned} \tag{7}$$

The outcome is $P(\boldsymbol{L}|do(\mathbf{H}=h), \boldsymbol{T}) = P(\boldsymbol{L}|do(\boldsymbol{H}=h))$, which means that the $\boldsymbol{L}$ function is d-separated from $\boldsymbol{T}$ when $\boldsymbol{H}$ is set to $do(\boldsymbol{H}=h)$.

Although the $\boldsymbol{L}$ common cause function is similar to the causal chain, it is computed differently according to $P(\boldsymbol{X}_i|Pa(\boldsymbol{X}_i))$ because $\boldsymbol{H}$ in the common cause graph has no parent while $\boldsymbol{H}$ in the causal chain graph has $\boldsymbol{T}$ as its parent.

APPENDIX D

Given the collider, $\boldsymbol{T} \rightarrow \boldsymbol{H} \leftarrow \boldsymbol{L}$, the $\boldsymbol{L}$ function can be expressed as:

$$\begin{aligned} P\left(\boldsymbol{L} \mid do(\boldsymbol{H}=h), \boldsymbol{T}\right) &= \frac{P(do(\boldsymbol{H}=h)|\boldsymbol{T}, \boldsymbol{L})P(\boldsymbol{T})P(\boldsymbol{L})}{P(do(\boldsymbol{H}=h)|\boldsymbol{T}, \boldsymbol{L})P(\boldsymbol{T})} \\ &= P(\boldsymbol{L}) \end{aligned} \tag{8}$$

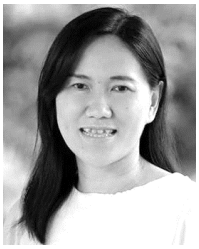
The outcome is $P(\boldsymbol{L}|do(\boldsymbol{H}=h), \boldsymbol{T}) = P(\boldsymbol{L})$, which means that the $\boldsymbol{L}$ function is d-separated to $\boldsymbol{T}$ even though $\boldsymbol{H}$ is set to $do(\boldsymbol{H}=h)$.

REFERENCES

[1] G. B. Berikol and G. Berikol, "Use of artificial intelligence in emergency medicine," in *Artificial Intelligence in Precision Health*, 2020, pp. 405–413.
[2] Y.-L. Chou, C. Moreira, P. Bruza, C. Ouyang, and J. Jorge, "Counterfactuals and causability in explainable artificial intelligence: Theory, algorithms, and applications," *Inf. Fusion*, vol. 81, pp. 59–83, May 2022.
[3] B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020.
[4] F. Ofli, F. Alam, and M. Imran, "Analysis of social media data using multimodal deep learning for disaster response," Tech. Rep., Apr. 2020.
[5] A. Kumar, J. P. Singh, Y. K. Dwivedi, and N. P. Rana, "A deep multimodal neural network for informative Twitter content classification during emergencies," *Ann. Oper. Res.*, pp. 1–32, Jan. 2020.
[6] N. Formosa, M. Quddus, S. Ison, M. Abdel-Aty, and J. Yuan, "Predicting real-time traffic conflicts using deep learning," *Accident Anal. Prevention*, vol. 136, Mar. 2020, Art. no. 105429.
[7] M. Anbarasan, B. Muthu, C. B. Sivaparthipan, R. Sundarasekar, S. Kadry, S. Krishnamoorthy, D. J. Samuel R., and A. A. Dasel, "Detection of flood disaster system based on IoT, big data and convolutional deep neural network," *Comput. Commun.*, vol. 150, pp. 150–157, Jan. 2020.
[8] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Mach. Intell.*, vol. 1, no. 5, pp. 206–215, 2019.
[9] T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," *Artif. Intell.*, vol. 267, pp. 1–38, Feb. 2019.
[10] D. Ben-Israel, W. B. Jacobs, S. Casha, S. Lang, W. H. A. Ryu, M. de Lotbiniere-Bassett, and D. W. Cadotte, "The impact of machine learning on patient care: A systematic review," *Artif. Intell. Med.*, vol. 103, Mar. 2020, Art. no. 101785.
[11] C. Son, F. Sasangohar, T. Neville, S. C. Peres, and J. Moon, "Investigating resilience in emergency management: An integrative review of literature," *Appl. Ergonom.*, vol. 87, Sep. 2020, Art. no. 103114.
[12] D. Reynard and M. Shirgaokar, "Harnessing the power of machine learning: Can Twitter data be useful in guiding resource allocation decisions during a natural disaster?" *Transp. Res. D, Transp. Environ.*, vol. 77, pp. 449–463, Dec. 2019.
[13] A. Devaraj, D. Murthy, and A. Dontula, "Machine-learning methods for identifying social media-based requests for urgent help during hurricanes," *Int. J. Disaster Risk Reduction*, vol. 51, Dec. 2020, Art. no. 101757.
[14] S. Madichetty, "Identification of medical resource tweets using majority voting-based ensemble during disaster," *Social Netw. Anal. Mining*, vol. 10, no. 1, p. 66, Dec. 2020.
[15] S. Sarkar, A. Pramanik, J. Maiti, and G. Reniers, "Predicting and analyzing injury severity: A machine learning-based approach using class-imbalanced proactive and reactive data," *Saf. Sci.*, vol. 125, May 2020, Art. no. 104616.
[16] X. Yu, C. Li, W.-X. Zhao, and H. Chen, "A novel case adaptation method based on differential evolution algorithm for disaster emergency," *Appl. Soft Comput.*, vol. 92, Jul. 2020, Art. no. 106306.
[17] Y.-H. Kuo, N. B. Chan, J. M. Y. Leung, H. Meng, A. M.-C. So, K. K. F. Tsoi, and C. A. Graham, "An integrated approach of machine learning and systems thinking for waiting time prediction in an emergency department," *Int. J. Med. Informat.*, vol. 139, Jul. 2020, Art. no. 104143.

[18] R. Costache, M. C. Popa, D. Tien Bui, D. C. Diaconu, N. Ciubotaru, G. Minea, and Q. B. Pham, "Spatial predicting of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning," *J. Hydrol.*, vol. 585, Jun. 2020, Art. no. 124808.
[19] X. Zhao, R. Lovreglio, and D. Nilsson, "Modelling and interpreting pre-evacuation decision-making using machine learning," *Autom. Construct.*, vol. 113, May 2020, Art. no. 103140.
[20] M. Yu, M. Bambacus, G. Cervone, K. Clarke, D. Duffy, Q. Huang, J. Li, W. Li, Z. Li, Q. Liu, and B. Resch, "Spatiotemporal event detection: A review," *Int. J. Digit. Earth*, vol. 13, no. 2, pp. 1339–1365, 2020.
[21] C. Albrecht, B. Elmegreen, O. Gunawan, H. F. Hamann, L. J. Klein, S. Lu, F. Mariano, C. Siebenschuh, and J. Schmude, "Next-generation geospatial-temporal information technologies for disaster management," *IBM J. Res. Dev.*, vol. 64, no. 1, pp. 1–12, 2020.
[22] B. Sahoh and A. Choksuriwong, "Automatic semantic description extraction from social big data for emergency management," *J. Syst. Sci. Syst. Eng.*, vol. 29, no. 4, pp. 412–428, Aug. 2020.
[23] M. A. Abebe, J. Tekli, F. Getahun, R. Chbeir, and G. Tekli, "Generic metadata representation framework for social-based event detection, description, and linkage," *Knowl.-Based Syst.*, vol. 188, Jan. 2020, Art. no. 104817.
[24] Z. Xu, Y. Liu, N. Y. Yen, L. Mei, X. Luo, X. Wei, and C. Hu, "Crowdsourcing based description of urban emergency events using social media big data," *IEEE Trans. Cloud Comput.*, vol. 8, no. 2, pp. 387–397, Apr. 2020.
[25] J. Pearl and D. Mackenzie, "The book of why: The new science of cause and effect," *Basic Books*, to be published.
[26] B. Sahoh and A. Choksuriwong, "Towards smart emergency management: Trends and challenges of feature engineering," in *Proc. 22nd Int. Comput. Sci. Eng. Conf. (ICSEC)*, Nov. 2018, pp. 1–4.
[27] D. J. Hilton and B. R. Slugoski, "Knowledge-based causal attribution: The abnormal conditions focus model," *Psychol. Rev.*, vol. 93, no. 1, pp. 75–88, 1986.
[28] J. Pearl, "The seven tools of causal inference, with reflections on machine learning," *Commun. ACM*, vol. 62, no. 3, pp. 54–60, Feb. 2019.
[29] G. R. Mayes, "Argument-explanation complementarity and the structure of informal reasoning," *Informal Log.*, vol. 30, no. 1, pp. 92–111, 2010.
[30] L. Michael and D. W. Eric-Jan, *Bayesian Cognitive Modeling: A Practical Course*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[31] J. Pearl, *Causality: Models, Reasoning and Inference*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[32] J. Pearl, "The do-calculus revisited," in *Proc. 28ht Conf. Uncertainty Artif. Intell.*, Aug. 2012, pp. 4–11.
[33] M. A. Sit, C. Koylu, and I. Demir, "Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: A case study of Hurricane Irma," *Int. J. Digit. Earth*, to be published.
[34] T. J. VanderWeele, M. A. Hernán, and J. M. Robins, "Causal directed acyclic graphs and the direction of unmeasured confounding bias," *Epidemiology*, vol. 19, no. 5, pp. 720–728, Sep. 2008.
[35] S. R. Cole, R. W. Platt, E. F. Schisterman, H. Chu, D. Westreich, D. Richardson, and C. Poole, "Illustrating bias due to conditioning on a collider," *Int. J. Epidemiol.*, vol. 39, no. 2, pp. 417–420, Apr. 2010.
[36] D. Barber, *Bayesian Reasoning and Machine Learning*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[37] T. Oroszi, "A preliminary analysis of high-stakes decision-making for crisis leadership," *J. Bus. Contin. Emer. Plan.*, vol. 11, no. 4, pp. 335–359, 2018.
[38] Ş. U. Zübeyde and G. Y. Ali, "The effects of globalization and terrorism on tourist arrivals to Turkey," in *Strategies in Sustainable Tourism, Economic Growth and Clean Energy*, 2021, pp. 109–123.
[39] S. Jitpiromsri, N. Waitoolkiat, and P. Chambers, "Special issue: Quagmire of violence in Thailand's southern borderlands chapter 1: Introduction," *Asian Affairs, Amer. Rev.*, vol. 45, no. 2, pp. 43–55, Apr. 2018.
[40] J. Wang, S. Ni, S. Shen, and S. Li, "Empirical study of crowd dynamic in public gathering places during a terrorist attack event," *Phys. A, Stat. Mech. Appl.*, vol. 523, pp. 1–9, Jun. 2019.
[41] R. O. Mujalli, G. López, and L. Garach, "Bayes classifiers for imbalanced traffic accidents datasets," *Accident Anal. Prevention*, vol. 88, pp. 37–51, Mar. 2016.
[42] S. Jitpiromsri, "Deep south watch database: Cumulative incidents in deep Southern Thailand January 2004- August2020," Tech. Rep., 2020.
[43] C. Loynes, J. Ouenniche, and J. De Smedt, "The detection and location estimation of disasters using Twitter and the identification of non-governmental organisations using crowdsourcing," *Ann. Oper. Res.*, vol. 308, pp. 339–371, Jul. 2020.
[44] J. Kersten and F. Klan, "What happens where during disasters? A workflow for the multifaceted characterization of crisis events based on Twitter data," *J. Contingencies Crisis Manage.*, vol. 28, no. 3, pp. 262–280, Sep. 2020.
[45] E. Bareinboim and J. Pearl, "Controlling selection bias in causal inference," *J. Mach. Learn. Res.*, vol. 22, pp. 100–108, 2012.
[46] K. Murphy, "An introduction to graphical models," *Rap. Tech.*, vol. 96, pp. 1–19, May 2001.

BUKHOREE SAHOH received the B.S. degree in computer science from the Prince of Songkla University (PSU), Songkhla, Thailand, in 2010, the M.Sc. degree in information technology from Thaksin University, Phatthalung, Thailand, in 2015, and the Ph.D. degree in computer engineering from PSU, in 2020. He is currently a Lecturer with the School of Informatics, Walailak University (WU), Nakhon Si Thammarat, Thailand, where he is a member of the Informatic Innovation Center of Excellence (IICE). His current research interests include natural language processing, complex event understanding, explainable artificial intelligence, and causality in machine learning.



KANJANA HARUEHANSAPONG received the B.B.A. degree in business computer from the Prince of Songkla University (PSU), Songkhla, Thailand, in 1994, and the M.Sc. degree in management of information technology from Walailak University (WU), Nakhon Si Thammarat, Thailand, in 2007. She is currently a Lecturer with the School of Informatics, Walailak University (WU). Her current research interests include data mining and database systems.

MALLIKA KLIANGKHLAO received the B.S. degree in information and communication technology from Silpakorn University, Thailand, in 2010, and the M.Sc. degree in information technology from Thaksin University, Thailand, in 2014. She is currently pursuing the Ph.D. degree in computer engineering with the Prince of Songkla University, Thailand. Her research interests include variety of different topics, including causal inference, explainable artificial intelligence, and modern supply chain management.

View publication stats