Causal Inference Using Bayesian Network for Search and Rescue

Amanda Belden
Committee: Dr. Franz Kurfess (CS), Dr. Hunter Glanz, Dr. Soma Roy

Overview

01 Background

02 Methods

03 Model Building

04 Causal Inference

05 Conclusion

01 Background

Project Background

AI for Search and Rescue (SAR)

- Research project in the CS department since 2021
- Under the direction of Dr. Franz Kurfess, CS and two SMEs
- Aims to build a platform and AI models to improve Search and Rescue mission outcomes in the US

Search Methodologies

Probabilities of subjects movements and survival based on their categorization

-> Aim to improve this method by having one model that represents all categories of subjects and adds nuance to probabilities based on other variables

Hiker

Distance (horizontal) from the IPP (miles)					
	Temperate		Dry		Urban
	Mtn	Flat	Mtn	Flat	
n	568	274	221	58	8
25%	0.7	0.4	1.0	0.8	
50%	1.9	1.1	2.0	1.3	1.6
75%	3.6	2.0	4.0	4.1	
95%	11.3	6.1	11.9	8.1	

Distance (horizontal) from the IPP (kilometers)					
	Temperate		Dry		Urban
	Mtn	Flat	Mtn	Flat	
n	568	274	221	58	8
25%	1.1	0.6	1.6	1.3	
50%	3.1	1.8	3.2	2.1	2.6
75%	5.8	3.2	6.5	6.6	
95%	18.3	9.9	19.3	13.1	

Elevation (vertical) Change from the IPP (feet)						
	Temperate			Dry		
	Uphill	Down	Same	Uphill	Down	Same
%	32%	52%	16%	48%	52%	
25%	182	160		317	500	
50%	480	400		956	975	
75%	1175	1166		1500	2109	
95%	2634	2175		3623	5094	

Horizontal Change from IPP (miles) for Mtn Terrain						
	Temperate			Dry		
	Uphill	Down	Same	Uphill	Down	Same
n	58	131	34	47	57	0
25%	0.5	0.7	0.0	1.8	1.0	
50%	1.4	1.7	0.0	2.2	2.0	
75%	2.6	4.0	1.5	4.0	5.0	
95%	7.2	17.4	12.8	10.7	12.3	

Hiker

	Mobility (hours)	
	Temperate	Dry
n	232	112
25%	0	4
50%	3	8
75%	6	12
95%	14	26

	Dispersion Angle (degrees)	
	Temperate	Dry
n	134	28
25%	2	20
50%	23	47
75%	64	124
95%	132	175

Find Location (%)			
	Temp	Dry	Urban
n	312	196	17
Structure	13%	10%	24%
Road	13%	17%	35%
Linear	25%	31%	18%
Drainage	12%	18%	6%
Water	8%	9%	12%
Brush	2%	2%	
Scrub	3%	3%	
Woods	7%	6%	
Field	14%	1%	6%
Rock	4%	2%	

Scenario (%)	
n	2242
Avalanche	
Criminal	
Despondent	
Evading	1%
Investigative	1%
Lost	68%
Medical	2%
Drowning	
Overdue	16%
Stranded	4%
Trauma	7%

Survivability		
	Wilderness	Urban
Uninjured	78%	59%
Injured	16%	24%
Fatality	6%	12%
No Trace		6%
Survivability	**Alive**	**n**
<24 hours	97%	2460
>24 hours	76%	361
>48 hours	60%	118
>72 hours	52%	51
>96 hours	49%	23

Track Offset (meters)	
n	40
25%	50
50%	100
75%	238
95%	424

Why is exploring computer based models for SAR important?

Employing ML models can discover ways to **improve SAR missions** by discovering previously unknown relationships could be found that could **reduce mission time or casualties.**
Reduce operation time by having a more accessible way to access previous mission probabilities by changing from flipbook to evidence based model.

Dataset Information

~3000 missions from New York state, 80+ variables

Missing Subject

- Gender
- Age
- Fitness level
- Last known activity

SAR Team

- Number of resources (dogs, helicopters, rangers)
- Notification time
- Number agencies involved

Location

- Last known & found locations
- Elevation
- Type of terrain

Weather

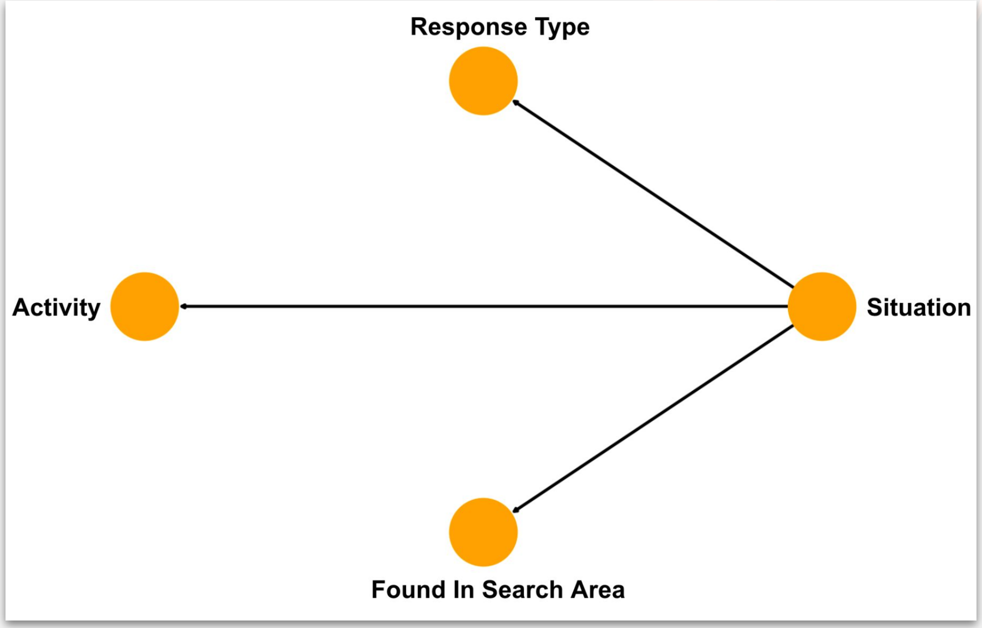
- Min/max temp
- Min/max snow depth
- Rain indicator

02 Methods

What is a Bayesian Network

A type of probability model that can be represented as a network:

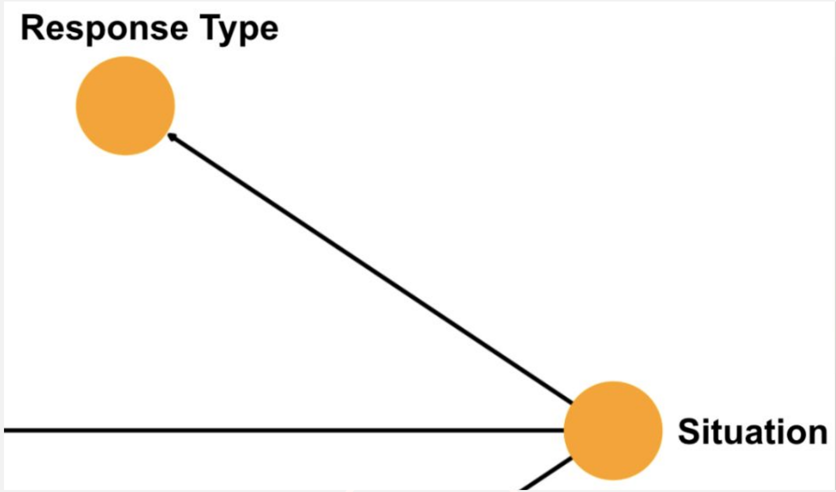
- each node is a variable
- each connecting arrow is an association between those variables



------------------- -------------------- ------------------- --------------------

Does Situation have an association or causal relationship with Response Type?

Association ≠ Causation



What is a Causal Inference?

- We can determine if an association is a causal relationship, where one variables causes a change in the probability of another.

- Experiments vs Observational Data

Do-Calculus and ACE

Based on our causal assumption that X is the cause variable and Y is the effect variable

- We can intervene on X, forcing it to be a certain value, denoted as $do(X=x)$.
- We can find both postintervention values of $P(Y=y|do[X=x])$ and $P(Y=y|do[X=x-])$
- Average Causal Effect (ACE) = $P(Y=y|do[X=x]) - P(Y=y|do[X=x-])$

-> We can make an inductive conclusion that $X = x$ is a probabilistic cause of $Y = y$ if the ACE is calculated to be positive, indicating a rise in the probability of $Y=y$ when X is changed from x- to x by our action

03 Model Building

Data Cleaning

- Data Aggregation
- Bin Quantitative Variables
- One-Hot Encoding
- Combine Like Categories and Variables
- Remove Sparse or Redundant Variables

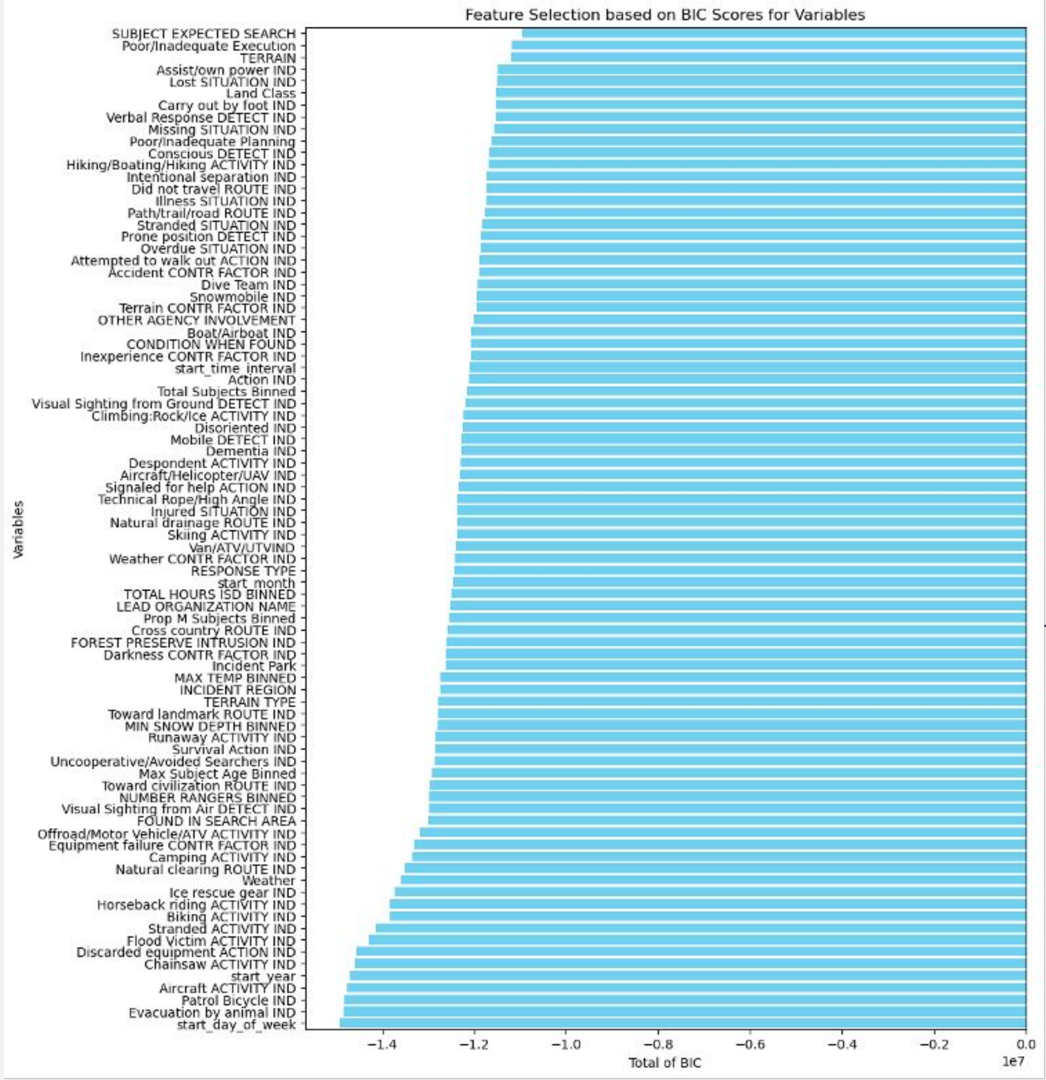
Model Specifications

- Uninformative prior -> assumes no associations
- Hill-Climb Search algorithm
  - Compares all one-step changes and selects best performing
- Uses BIC to compare models

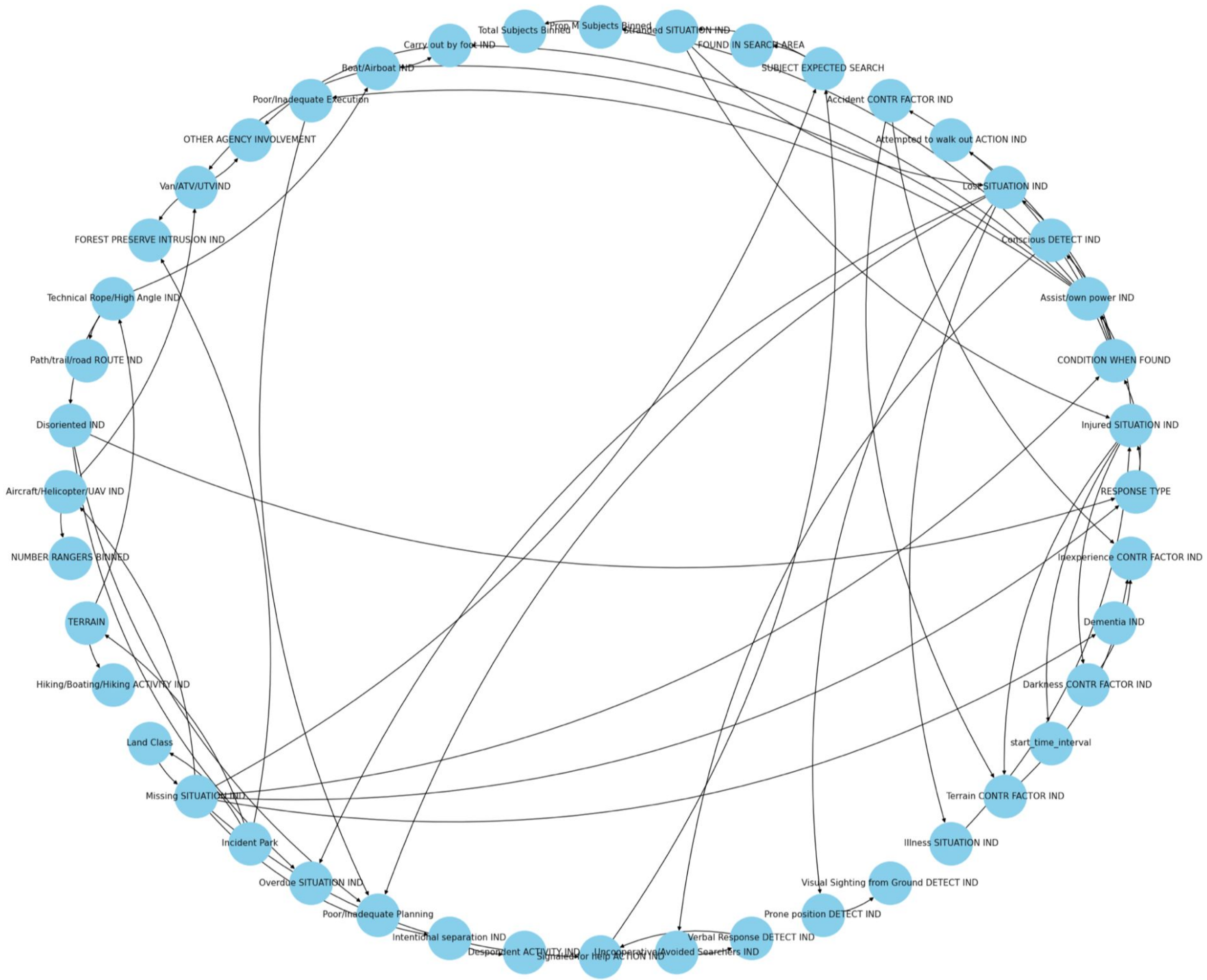
Modeling Process

1. Feature Selection Algorithm

2. Manual Pruning



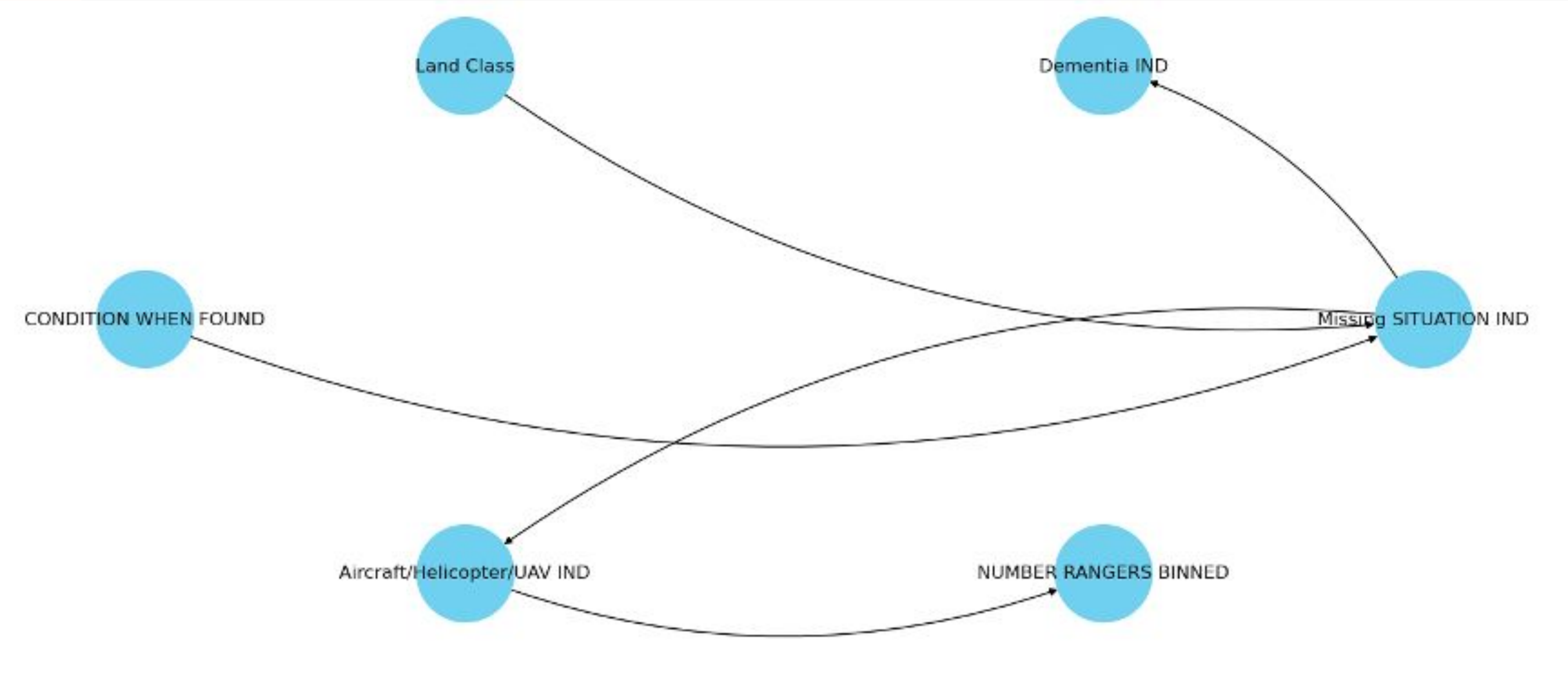
Final Model



Prop M Subjects Binned
Stranded SITUATION IND
FOUND IN SEARCH AREA
SUBJECT EXPECTED SEARCH
Accident CONTR FACTOR IND
Attempted to walk out ACTION IND
Lost SITUATION IND
Conscious DETECT IND
Assist/own power IND
CONDITION WHEN FOUND
Injured SITUATION IND
RESPONSE TYPE
Inexperience CONTR FACTOR IND
Dementia IND
Darkness CONTR FACTOR IND
start_time_interval
Terrain CONTR FACTOR IND
Illness SITUATION IND
Visual Sighting from Ground DETECT IND
Prone position DETECT IND
Verbal Response DETECT IND
Uncooperative/Avoided Searchers IND
Signaled for help ACTION IND
Despondent ACTIVITY IND
Intentional separation IND
Poor/Inadequate Planning
Overdue SITUATION IND
Incident Park
Missing SITUATION IND
Land Class
Hiking/Boating/Hiking ACTIVITY IND
TERRAIN
NUMBER RANGERS BINNED
Aircraft/Helicopter/UAV IND
Disoriented IND
Path/trail/road ROUTE IND
Technical Rope/High Angle IND
FOREST PRESERVE INTRUSION IND
Van/ATV/UTVIND
OTHER AGENCY INVOLVEMENT
Poor/Inadequate Execution
Boat/Airboat IND
Carry out by foot IND
Total Subjects Binned

04 Causal Inference

Example: Missing Indicator



Probability of Condition When Found Given do(Missing Situation)

do(Missing Situation Indicator)	Condition When Found	Probability
do(Yes)	Deceased	**0.3071**
	No Medical Assistance Required	0.3106
	Required Medical Treatment	0.3823
do(No)	Deceased	**0.0561**
	No Medical Assistance Required	0.5118
	Required Medical Treatment	0.4321

P(Deceased | do[Missing = Yes])
= 0.3071

P(Deceased | do[Missing = No])
= 0.0561

Average Causal Effect
= 0.3071 – 0.0561
= 0.251

Being missing *causes* a higher probability of being found deceased

05

Conclusion

Takeaways

- Bayesian networks can be used to represent the associations between factors that affect Search and Rescue missions
- Changing from a flipbook to a nuanced, evidence based model is possible
- Given more data they could discover previously unknown patterns between missions and between different types of missing subjects.
- As experienced SAR members age out and younger members mainly experience “easier” missions because of cell phones, it is important to codify as much knowledge as possible to be used in future missions in something more accessible and nuanced than a book

Future Works

- Compare informative priors from experienced SAR professionals
- Collate more data beyond NY
- Create interactive graph and interface, highlighting actionable nodes
- Applying other ML and DL models to SAR data

Questions?