Actuaries Institute.

Whole person. Whole system.

IDSS 2023

12 – 14 November
Hobart

Bring XAI to Claims Management

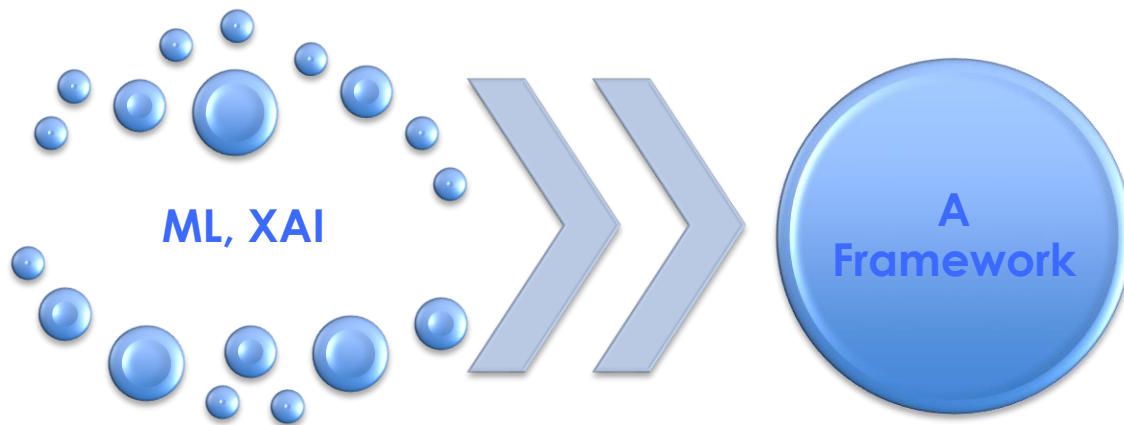## Jamie Wu

Paper Title

A Feasible Situation Awareness-Based Evaluation Framework for Introducing Explainable AI to Claims Management

Will be available to download/read on the event website.

Part 1. The aim

ML, XAI

A Framework

Provide powerful results and helpful explanations of the results

Evaluate whether the explanations are effective

We know ML can be powerful

Practical Problem to Solve	ML Method
What are the main factors affecting return-to-work after musculoskeletal injury?	Random forest, support vector machine, logistic regression
Identify occupational injuries in agriculture using existing administrative data.	Naïve Bayes machine learning
Predict the post-incident state of an injury and identify its contributory factors.	Multiple methods

We know ML can be black box

- A more academic way to say it is:
- We face the challenge of explainability when we hope to use a black box, knowing that it provides good performance (usually measured by accuracy).

- We have a perfect model, but we don't know how to explain the results, and/or we don't how the results are produced.

We hope to unbox it

- eXplainable Artificial Intelligence(XAI) or Interpretable Machine Learning are the flashlights.

- We will give examples of XAI techniques and explain more later.

- For now we just need to know that XAI provide explanations of how the black box works, and how the results should be interpreted.

- To some extent, we can say XAI helps us unbox the black box.

# We are required to explain it

1. Customer will ask questions

2. Stakeholders need to understand

3. Regulatory requirements

We need a framework, just like AAF

- AAF: Actuarial Advice Framework
- We need an evaluating framework to guide us
1. To develop sufficient explanation for the use of ML
2. To evaluate the effectiveness of explanation in given scenario
3. To make all stakeholders feel comfortable to accept the results produced by ML method

Part 2. What are XAI

lime
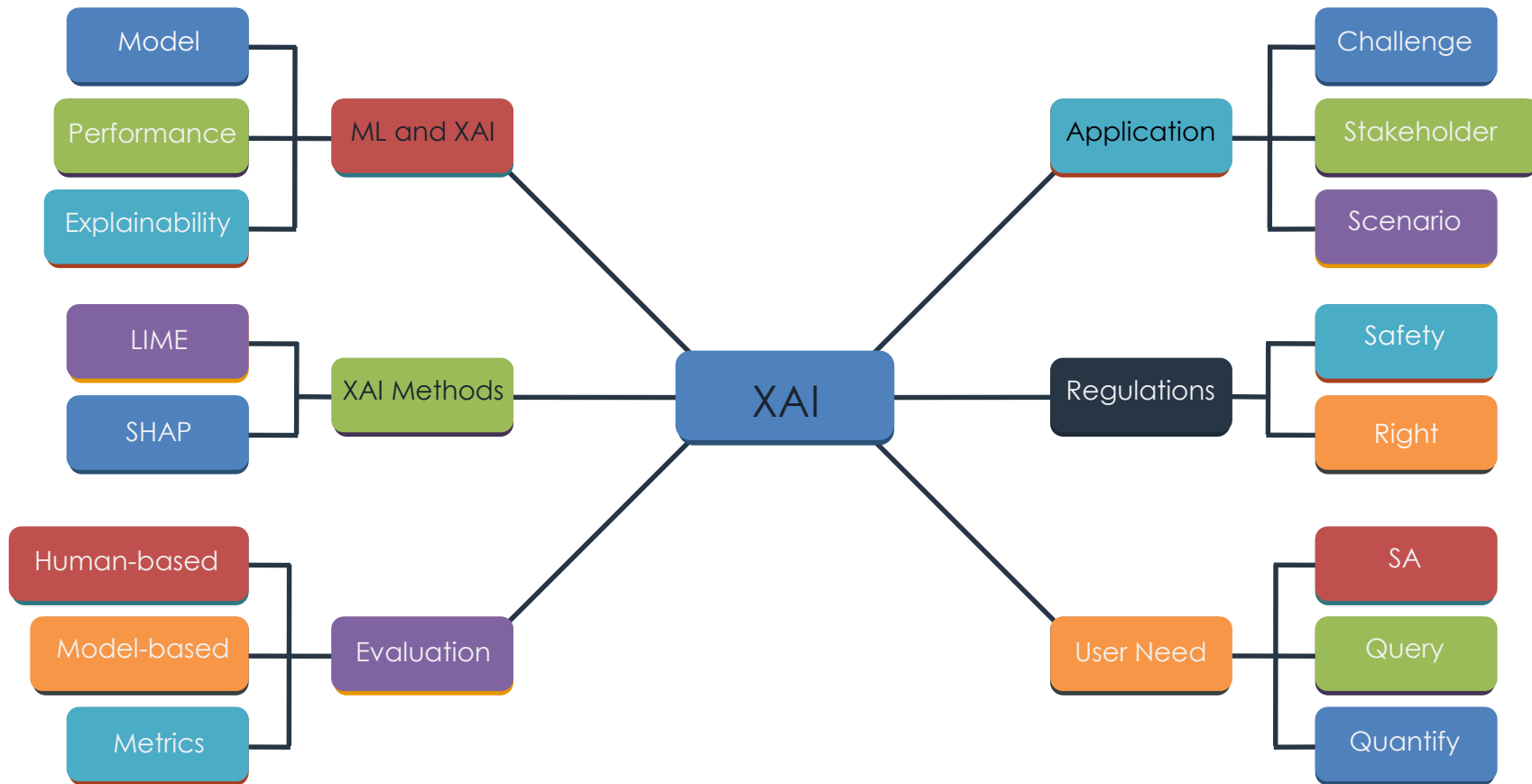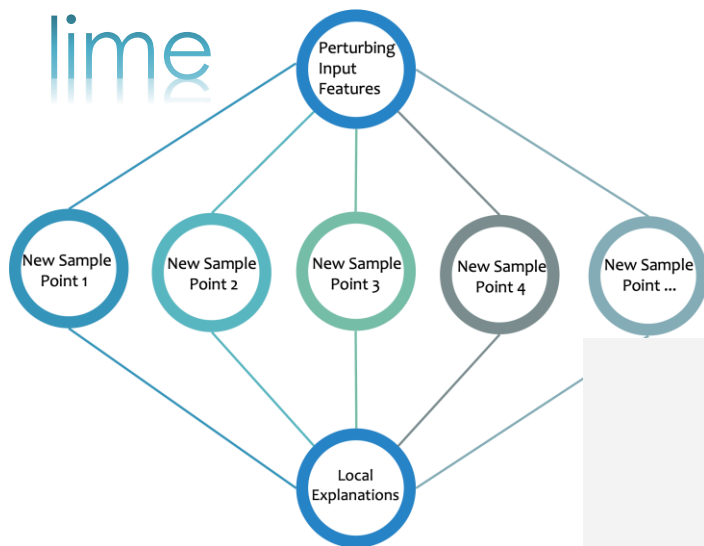
There are more
to explore...

SHAP

Python code to produce the SHAP result is provided on the next slide, the full illustrative case is available in the paper.

The provided code can be run in the environment of Python 3.
The output of the code is the summary plot of SHAP values on previous slide.

IDSS 2023

Actuaries
Institute.

```python
import shap
import numpy as np
import pandas as pd
from sklearn.linear_model import LogisticRegression
```

```python
# Sample dataset
data = pd.DataFrame({
    'Hours Spent on Recovering Plan Weekly': [10, 25, 5, 15, 30],
    'Number of Daily Medicine Taken': [2, 3, 1, 2, 4],
    'Returned to work': [0, 1, 0, 1, 1]  # 0 for not returned after a period of time, 1 for returned
})

X = data[['Hours Spent on Recovering Plan Weekly', 'Number of Daily Medicine Taken']]
y = data['Returned to work']

# Train a logistic regression model
model = LogisticRegression()
model.fit(X, y)


# Generate SHAP values for multiple instances
shap_values = explainer.shap_values(X)

# Create a summary plot for all instances
shap.summary_plot(shap_values, X)
```

Part 3. The evaluation framework

- In real life, we first start from a goal that we need to achieve.
- We hope to provide explanations that indeed help achieve the goal.
- We set a general goal as an illustrative example: **we hope to use the result of a machine learning method to understand the main drivers of injuries and disability claims.**
- This framework can be tailored to a specific scenario.

Claims Data in Injuries and Disability

- Claims data in injuries and disability usually includes not only basic demographic details, but also detailed information about individual's heath condition, such as the specifics about the nature of the injury or disability and the severity of the injury, the specific part of body that is affected, etc. If the injury or disability is work-related, the employer's detail and work condition will be in the data for assessment of eligibility of worker's compensation.

- The potential application of advanced machine learning methods is usually in the area of experience analysis for worker's compensation, so that more realistic assumptions on claims development can be made. As for the fraud claims management, we hope more accurate prediction of claims fraud can be performed, so that corresponding strategic adjustment on existing ratings can be adopted.

The problem of explainability comes from two sources.

One is that the ML technique itself could be technically demanding, and the interpretation of results requires deeper knowledge. People may feel discouraged to use difficult method because they don't have confidence and time to understand it.

Another problem is that a more advanced ML method can be a black box, and there is no simple answer of how it produces the results regardless of people's technical background.

In this context, we want to know whether the explanations produced by XAI can mitigate the challenge effectively.

Goal-Directed Task Analysis

- The study of Goal-Directed Task Analysis was performed by D. Jones and M. Endsley of SA Technologies, Inc. In GDTA(Goal-Directed Task Analysis), each goal would be broken down into sub-goals, decisions, and ultimately informational need for the decision-making in a given scenario.

- We proposed a GDTA interview work flow to determine the informational need in a given scenario.
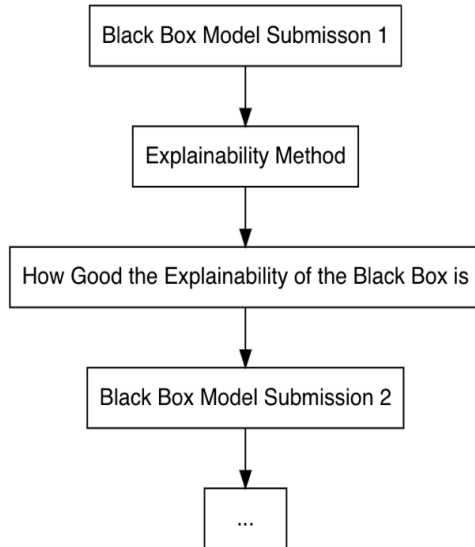
| Introduce GDTA in context | → | Questions and Open Discussions | → | Discussion on Decision Making |

The expected result of the interview is informational needs.
Explanations produced by XAI should meet those needs.
In the paper, we proposed to use Situation Awareness raised by Endsley
as a guidance to determine different levels of informational need.

User-based Evaluation Using Situation Awareness Metrics

Iterations to achieve the goal

Black Box Model Submisson 1

↓

Explainability Method

↓

How Good the Explainability of the Black Box is

↓

Black Box Model Submission 2

↓

...

The evaluation

- The central interest is to evaluate the effectiveness of explanations provided by XAI for a machine learning model.
- We are interested in whether the explanations provided by XAI method improve users' perceptions comparing to existing explanations that would be used if XAI tool is absent.
- We propose to use Situation-Awareness based metrics.
- The informational needs can be catigorised into different levels of requirements of situational awareness: perception (L1), comprehension (L2) and projection (L3)
- Experiment design to test the hypothesis can be found in the paper.

Thank you

IDSS 2023

12 – 14 November
Hobart

Actuaries
Institute.