ADVANCECRT

# Explainable AI in Healthcare

Urja Pawar, Dr. Ruairi O'Reilly, Dr. Christian Beder, Dr. Sean McSweeney, and Dr. Donna O'Shea
Munster Technological University (MTU)
Urja.Pawar@mycit.ie, Ruairi.Oreilly@mtu.ie, Christian.Beder@mtu.ie, Sean.McSweeney@mtu.ie, and Donna.Oshea@mtu.ie

What is eXplainable AI (XAI) ? - Techniques to interpret and explain AI-based systems

Why XAI?
- Understanding why an outcome was generated
- Explaining operation of black-box models
- Understanding the underlying training data
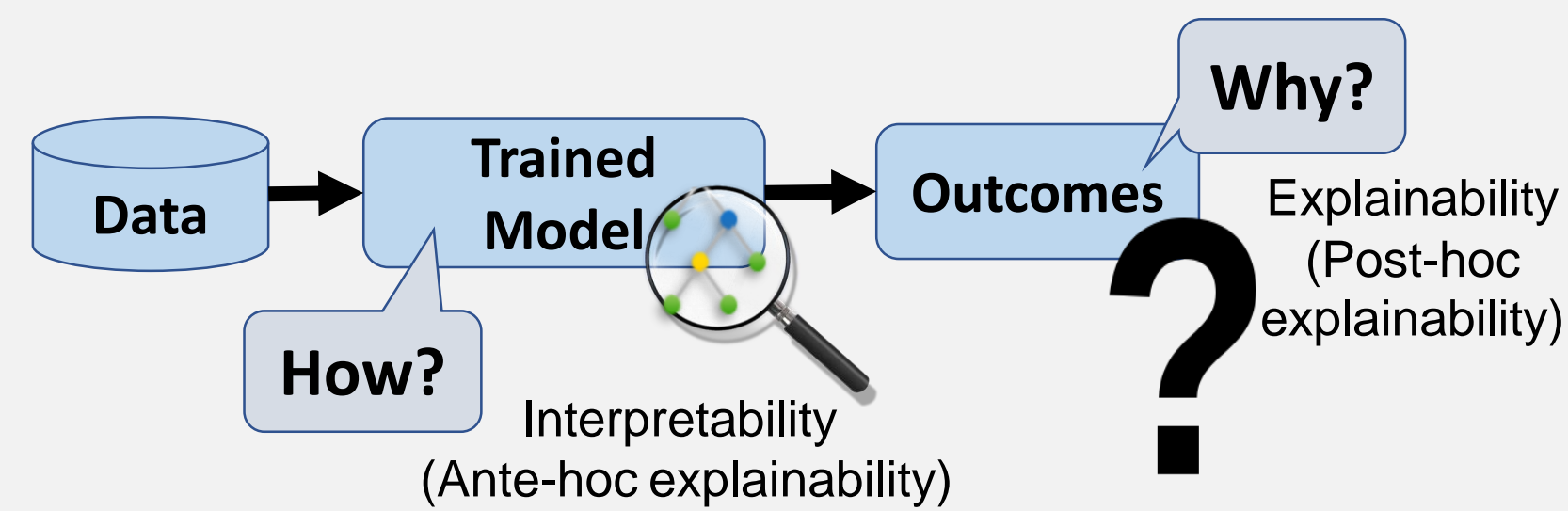
Feature Importance & Counterfactuals: Two of the commonly used XAI techniques for explaining machine learning based systems.

Feature Importance
Feature importance techniques assign scores to participating features based on their impact on model's outcome

Feature Importance (%)
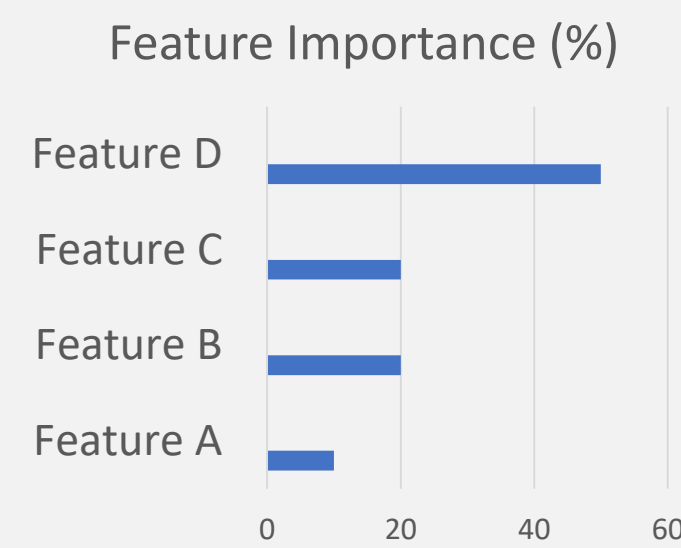
Counterfactuals
Counterfactual explanations are minimum changes in input data that will lead to a different classification

Test A	Test B
x	y

Classifier → 1

0: Healthy
1: Disease

Disease classification (binary)
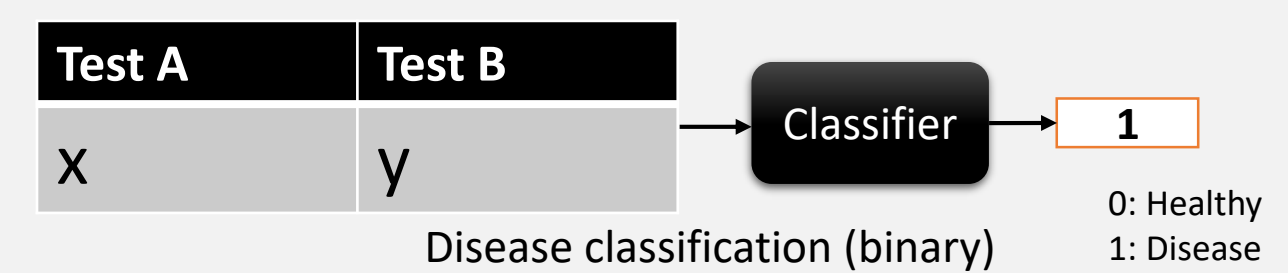
Test A	Test B	Output
-(no change)	t	0
z	-(no change)	0

Counterfactual explanations

About My Research
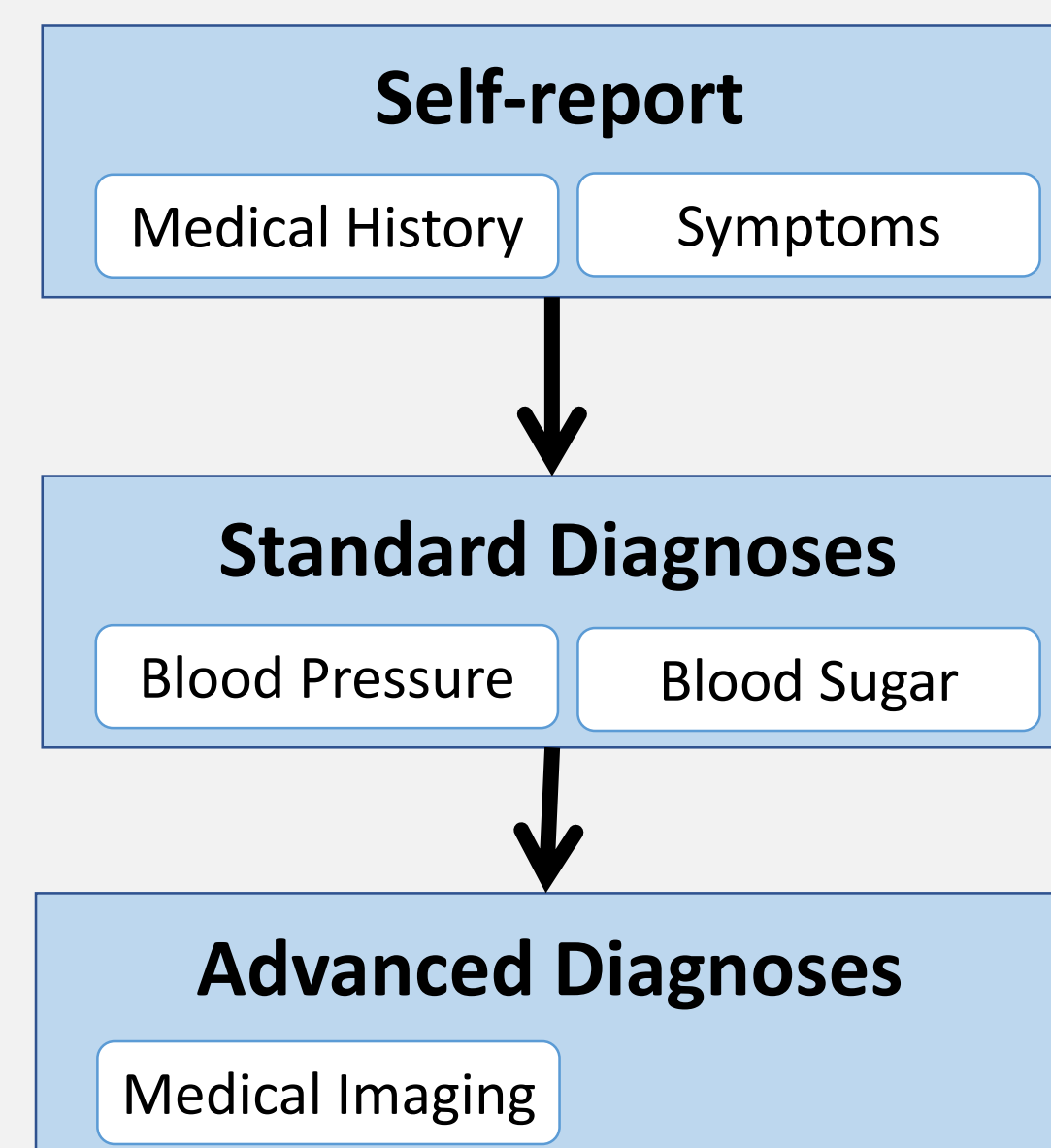
Generating Better Explanations

In Healthcare, knowledge of performance gains associated with diagnostic tests can enable easy validation

Along with feature importance, it is also essential to know the sensitivity of diagnostic tests as per the machine learning model

Information pertaining to underlying data is also important to highlight to enable transparency

Application Area: Hierarchical Medical workflows

- The healthcare system is hierarchical – small clinics for general check ups and hospitals for more advanced examinations
- AI-based systems should consider the hierarchical structure and the dynamic nature of medical workflows
- Current work focuses on enabling explainability in diagnostic workflows across the hierarchical medical settings in terms of performance gains, sensitivity of tests, and highlighting information about the underlying data utilised to train the machine learning model

Self-report
Medical History | Symptoms

↓

Standard Diagnoses
Blood Pressure | Blood Sugar

↓

Advanced Diagnoses
Medical Imaging

How can this project help solve the UN SDG challenges?

Goal 3: Good Health and Well-being

With XAI, there are better chances to:
- Trust and understand AI-based systems
- Adopt AI-based systems in clinical settings for early warning

These relate to goal of strengthening capacity of healthcare for early warning, and risk reduction of diagnostics.

Goal 9: Industry, Innovation, and Infrastructure

The more end-users understand AI-based systems, the better will be the adoption rate to use these systems

These will boost the technological and scientific advancements across the industries that can leverage AI-based systems to automate tasks and progress rapidly.

Goal 16: Peace, Justice And Strong Institutions

XAI is a sub-domain of Responsible AI that includes other factors while building AI-based systems such as data privacy, and bias.

Explainability of AI-based systems can be used for building more accountable and effective systems that are utilised in critical domains such as healthcare and finance.