XAI WORKSHOP

Introduction to XAI : Key Concepts and Definitions

IEEE SB ENICarthage

Achref Ben Ammar Machine Learning Engineer @Callem.ai

achref.benammar@ieee.org

WHAT IS XAI?

Processes and methods that make machine learning models interpretable. with a goal to help users understand and trust AI decisions.

Why is XAI Important?

- Trust and Transparency: Build user confidence in AI systems.
- Regulatory Compliance: Required by laws such as GDPR.
- Debugging & Improvement: Enables model enhancement by understanding decision-making.

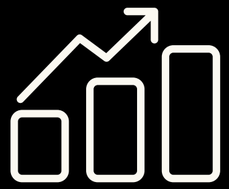
XAI IN CYBERSECURITY

Risk Mitigation

XAI helps explain why files or behaviors are flagged as malicious, making threat detection more reliable by reducing false positives or negatives.

Justification

Security teams can provide clear explanations for AI-driven decisions, ensuring that flagged threats or anomalies are understood and justified to stakeholders.

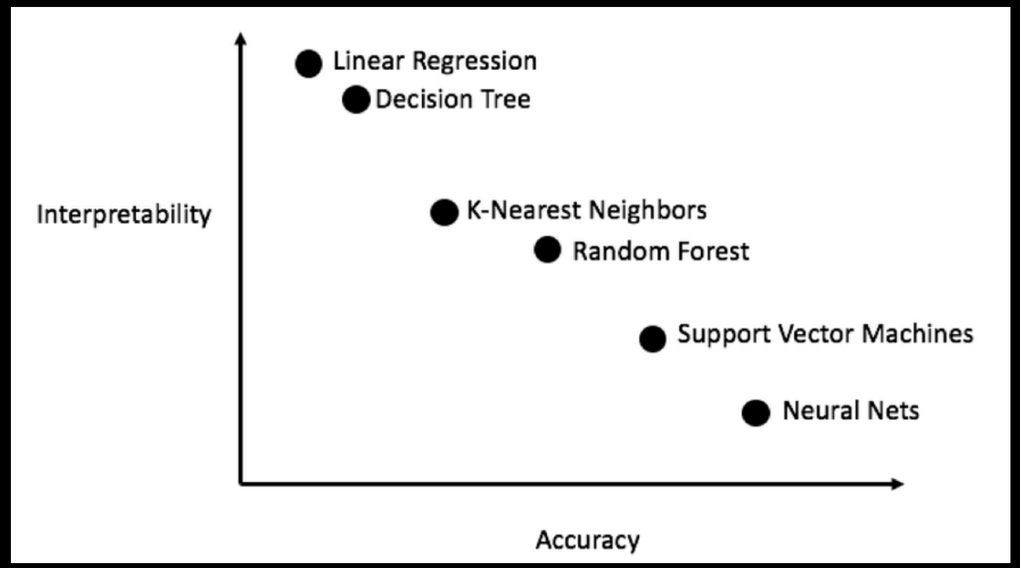


Improving Model Robustness:

By exposing decision processes, XAI helps identify weaknesses or biases in models, allowing teams to improve and harden cybersecurity defenses.

EXPLAINABILITY VS PERFORMANCE

This plot visually demonstrates how more interpretable models like Linear Regression and Decision Trees typically offer lower accuracy on complex tasks, while models with higher accuracy like Neural Networks and Support Vector Machines sacrifice interpretability, making them harder to explain.

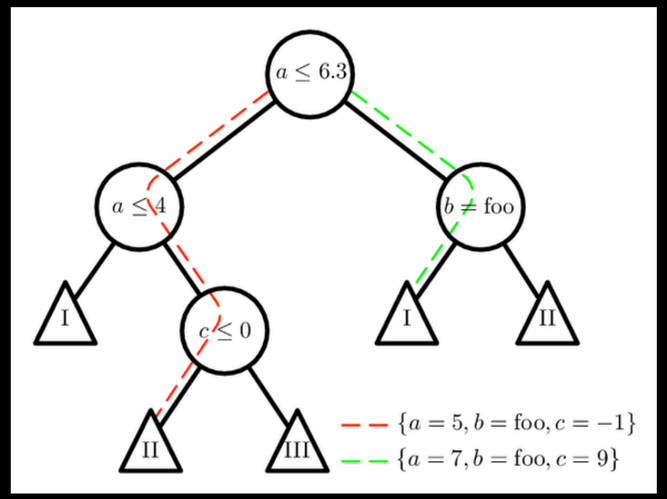


DECISION TREES FOR MALWARE DETECTION

Definition:
A decision tree is a supervised machine learning algorithm used for both classification and regression tasks.

How does it work?

The tree partitions data into subsets by making decisions based on feature values. Each internal node corresponds to a decision (based on a feature), and each leaf node corresponds to a predicted class or value.



LET’S APPLY

Explore Explainable AI with Decision Trees

Now that we've discussed Explainable AI (XAI) and its importance, it's time to put theory into practice.

We will dive into using Decision Trees, one of the most interpretable machine learning models, to understand how it helps in making AI decisions transparent and understandable.

Together, we’ll build a Decision Tree to classify malware and explore how and why decisions are made at each step.