TMRYK MODEL RISK MANAGEMENT

Why Tmryk?

TMRYK (Threat Management for Risks You Know) – your frontline defense against adversarial AI threats, ensures comprehensive security tailored to the unique demands of the enterprise AI ecosystem.

Redefining AI Security

Our platform acknowledges the challenges of integrating internal and third-party AI/ML models developed by data scientists.

Benefits

Understanding Model Risks

Understand the depth and breadth of AI vulnerabilities in your enterprise AI/ML Models

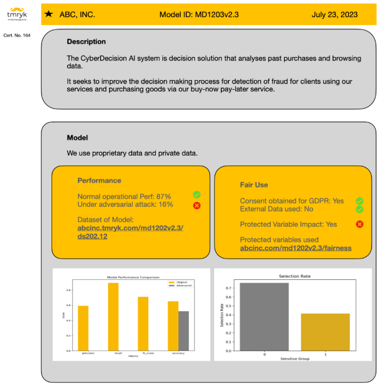
Prioritized Risk Mitigation:

Strategically address the most pressing threats, enhancing overall security posture.

Boost Compliance:

With TMRYK, remain ahead in the compliance curve and safeguard your business assets.

AI Model Repport



★ ABC, INC. Model ID: MD1203v2.3 July 23, 2023

Cert. No. 164

Description

The CyberDecision AI system is decision solution that analyses past purchases and browsing data.

It seeks to improve the decision making process for detection of fraud for clients using our services and purchasing goods via our buy-now pay-later service.

Model

We use proprietary data and private data.

Performance

Normal operational Perf: 87% ✓
Under adversarial attack: 16% ☒

Dataset of Model:
abcinc.tmryk.com/md1202v2.3/ds202.12

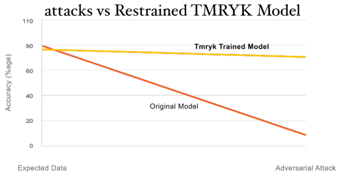
Fair Use

Consent obtained for GDPR: Yes ✓
External Data used: No ✓

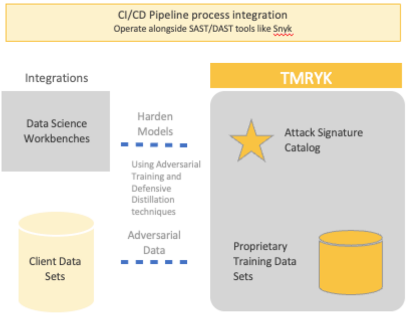
Protected Variable Impact: Yes ☒

Protected variables used
abcinc.com/md1202v2.3/fairness

Model Performance under adversarial AI attacks vs Restrained TMRYK Model



TMRYK Model Risk Management Service


CI/CD Pipeline process integration
Operate alongside SAST/DAST tools like Snyk
Integrations
Data Science Workbenches
Client Data Sets
Harden Models
Using Adversarial Training and Defensive Distillation techniques
Adversarial Data
TMRYK
Attack Signature Catalog
Proprietary Training Data Sets

APIs support most common AI/ML model types likes Decision Trees, XGBoost, Logistic Regression and Neural Networks

Library for execution via the Notebooks used by data scientists also available

Model Risk Management Service

Run Adversarial AI Attacks against models to test model robustness against possible Data Leakage, Model Extraction or Model Evasion cyberattacks

Generate synthetic adversarial datasets to retrain models for corresponding MITRE ATLAS attacks

Use the generated PDFs for AI solution compliance for ISO27001 and upcoming AI regulations

Integrate your models and datasets with our service to capture Model Performance

Automatically generate model feature explainability and fairness metrics

Validate 3rd model risk as vendor due diligence