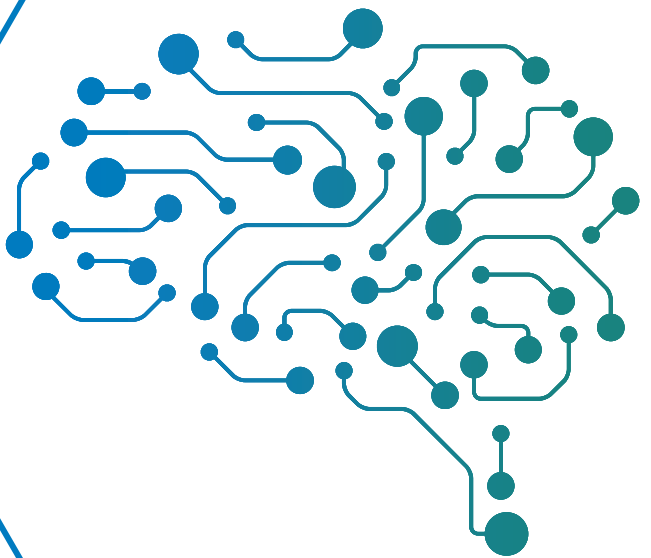
AIShield
Powered by Bosch

CASE STUDIES

Securing AI systems of the world across industries and use cases



Use Case | Automated Visual Inspection

Protecting Smart Manufacturer's Intellectual Property

Algorithm - Automated Visual Inspection

Industry – Manufacturing / Industry 4.0

Customer – Europe based Automotive parts manufacturer

AI Threat Prevented – Model Extraction, Evasion & Poisoning

System

Design a System. which can handle a variety of complex products

Image capture module Capture

Design a general-purpose image capture module for detection / product inspection for different kind of defects

Algorithms and frameworks

Provide algorithms and frameworks so that users can train defect detectors

Background & Objective

- For the manufacturer, product quality control is essential in order to fulfill the highest quality demands of customers. Automated production lines accomplish this in part with highly customized optical inspection systems. Systems comprised cameras & automated inspection algorithm to capture images and detecting defects on products
- Extraction of this highly adaptable, task-flexible and reconfigurable algorithm can enable malpractitioners to steal the intellectual property of the smart manufacturer who invested considerable time and effort in building and training the high accuracy algorithm

Solution Highlights

- Identified & demonstrated AI security threats to automated visual inspection algorithm by performing threat & risk analysis for AI model and system
- Developed a defense model engineered for AI/ML model and Integrated it the target smart factory environment

Result

Prevented loss of IP of an automated visual inspection algorithm with impact value of **USD 5 million.**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield
Powered by Bosch

Use Case | COVID Detection from X-ray Scans

Proving AI vulnerability of a healthcare startup

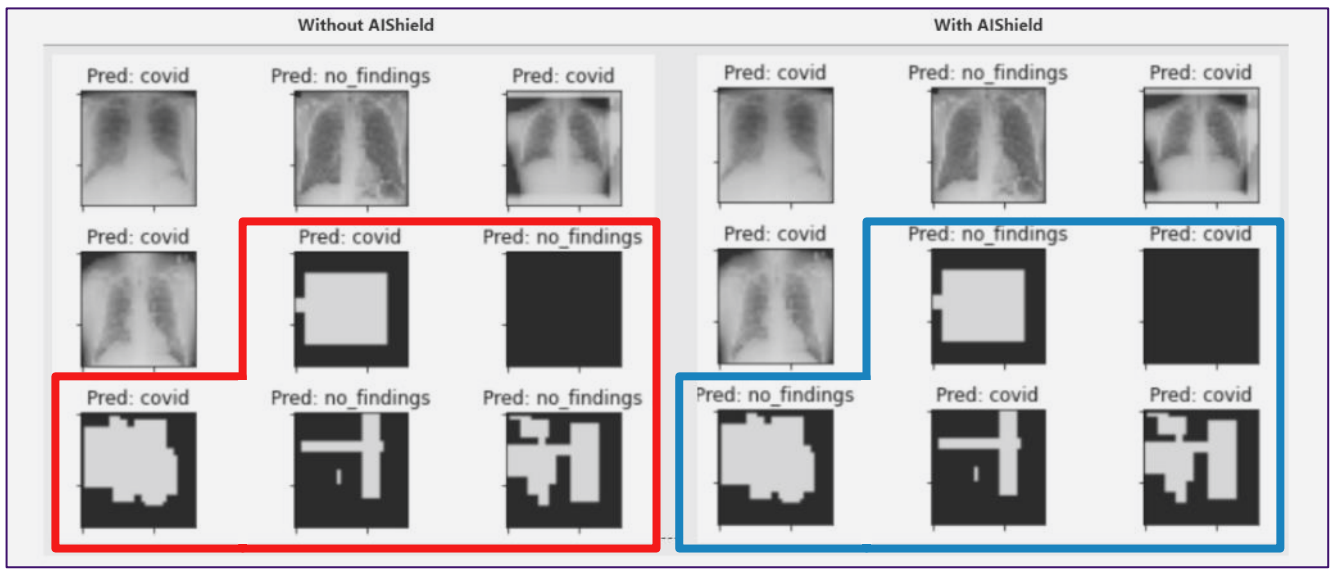
Algorithm – Automatic Detection Method for COVID from X-Ray

Industry – Healthcare

Customer – Healthcare startup

AI Threat Prevented – Model Extraction & Model Evasion

Visual Report



Background & Objective

A healthcare company using an AI algorithm to use X ray scans to make COVID decisions wanted to protect its algorithm from any theft or external manipulation. The algorithms was developed from multiple R&D studies and lots of valuable datasets.

Solution Highlights

AIShield's proprietary AI Security module implementation for:

- AI application risk assessment during production
- AI Security real time monitoring post deployment

Demonstration of protection of the AI application involving:

- Efficacy of defense model
- Performance metric (i.e., accuracy) of original and shielded model

Result

Increased patient safety and demonstrated the prevention of loss of intellectual property & revenue for the healthcare company with **impact value of USD 10+ million.**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield
Powered by Bosch

Use Case | Cancer Cell Detection

Proving AI vulnerability of a healthcare leader

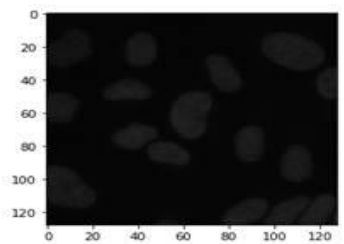
Algorithm – Automatic Detection Method for Cancer Cell Nucleus

Industry – Healthcare

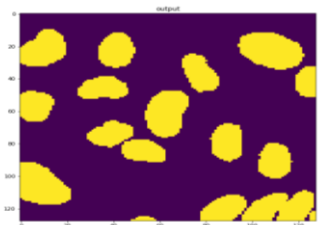
Customer – Multinational Science and Technology company

AI Threat Prevented – Model Extraction

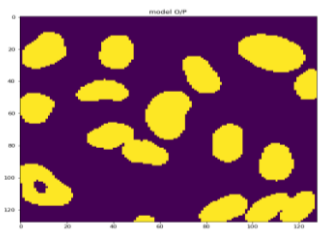
01 Original Input



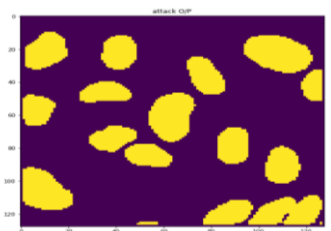
02 Human labeled Ground Truth



03 Original Model Output



04 Stolen Model Output



Background & Objective

The healthcare company was working on study that the mitotic defects and micronuclei in cancer cells can be used as biomarkers to evaluate the instability of the chromosomes. The study further aimed to detect cells with mitotic defects and micronuclei by applying an approach integrating the application of a convolutional neural network for normal cell identification and the proposed color layer signature analysis (CLSA) to spot cells with mitotic defects and micronuclei.

Solution Highlights

- AIShield team extracted the CNN algorithm developed over 6 months within 2 hours & only 8% delta to original model accuracy, as part of an ethical hacking case
- As part of threat mitigation AIShield team integrated a defense model with the original model which brought down the stolen model accuracy to only 10%

Result

Demonstrated the prevention of loss of intellectual property & revenue for the healthcare company with **impact value of USD 5 million.**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield Powered by Bosch

Use Case | Automated Loan Processing

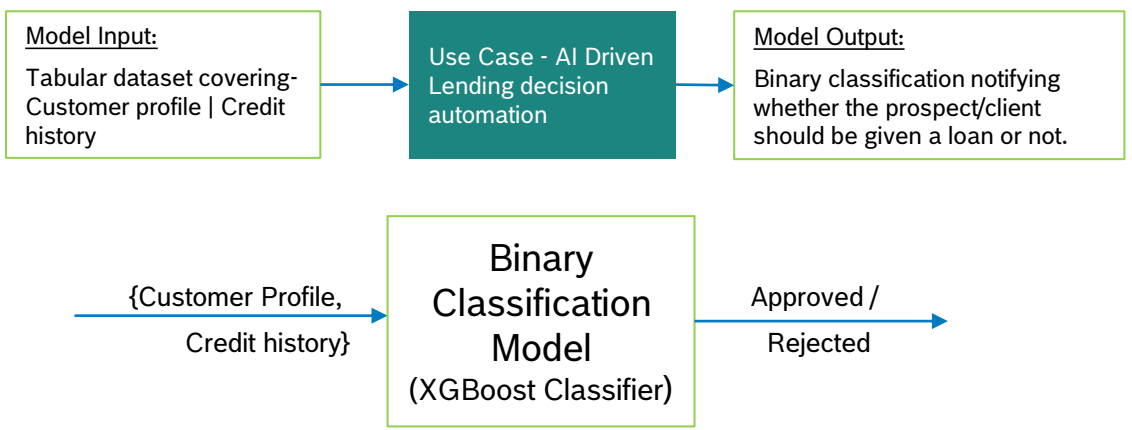
Reducing operational losses from fraud

Algorithm – AI Driven Lending Automation

Industry – BFSI

Customer – A large Banking MNC

AI Threat Prevented – Model Evasion



Background & Objective

A prominent UK bank, handling over 30% of the nation's £300 billion debit and credit card transactions, experienced nearly £70 million in external fraud losses in 2021 despite having a fraud detection model. This represented about 70% of total losses, marking it as the highest operational risk. The bank's CTO—along with AI/ML, risk and cybersecurity teams—aim to strengthen the current fraud detection model to decrease these losses and increase profit margins.

Solution Highlights

AIShield's proprietary AI Security module implementation for:

- AI application risk assessment during production
- AI Security real time monitoring post deployment

Demonstration of protection of the AI application involving:

- Efficacy of defense model
- Performance metric (i.e., accuracy) of original and shielded model

Result

With its transparent and objective assessment of residual risk, the solution has the potential to significantly improve the bank's ability to reduce residual operational risk by up to 15%. The risk management team is now able to make quick and secure releases of fraud detection models in just eight hours, 11x faster and with a 30x productivity boost compared to the previous process.

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield
Powered by Bosch

Use Case | Premium Services and CLV Prediction

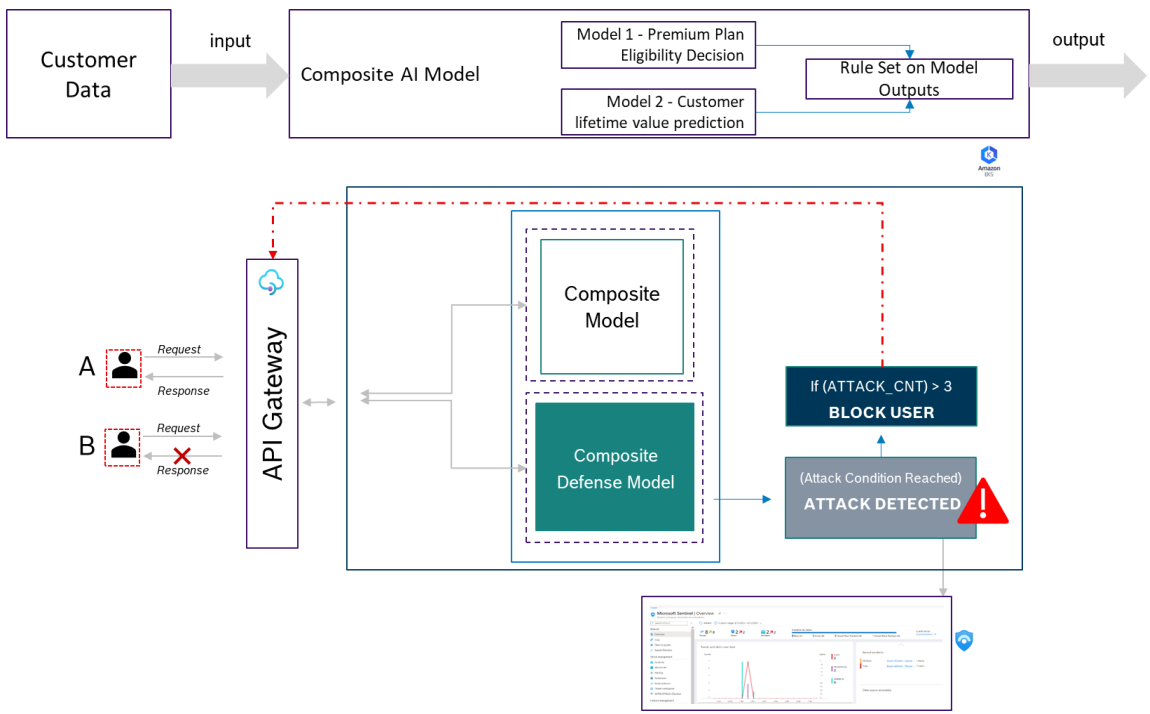
Reducing operational losses from fraud

Algorithm – AI Driven Premium Services Eligibility & CLV prediction

Industry – Telecommunications

Customer – Telecommunications System Integrator

AI Threat Prevented – Model Evasion



Background & Objective

Telecommunication industry use case: Multi-task AI model with two outputs (Binary Classification, Regression).

1. Approve/Reject request for premium services on credit through AI automation or virtual assistants. Approve/Reject request for premium services on credit through AI automation or virtual assistants.
2. Customer Lifetime Value (CLV) prediction. Failing to predict this value may result in profit loss for the telecommunication company. Threats could be from competitor, insider threat, or any user with malicious intent to cause financial damage.

Solution Highlights

Any potential manipulation to the AI system or any potential input to extract model knowhow with the intention of stealing will be flagged by the “secure” AI system.

Result

Reduction in operational risks and improved profitability.

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield Powered by Bosch

Use Case | Pedestrian detection algorithm

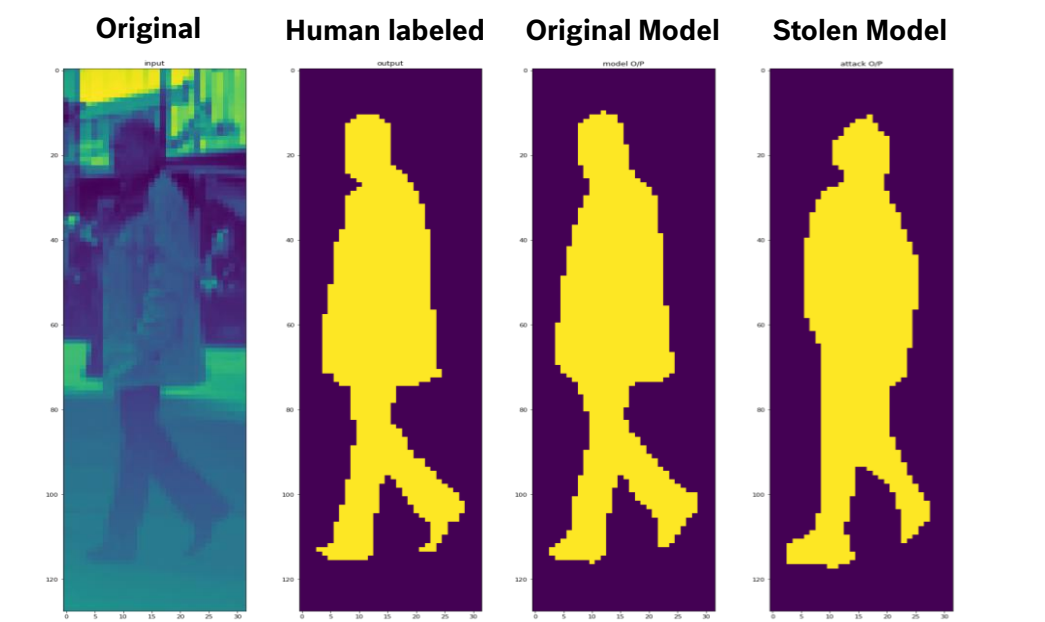
Preventing IP theft and ensuring safety

Algorithm – Pedestrian Detection

Industry – Automotive (Autonomous Driving)

Customer – Automotive company

AI Threat Prevented – Model Extraction



Background & Objective

Recover a functionally equivalent model by iteratively querying the model. This allows an attacker to examine the offline copy of the model and re-use the same without license or launch further attacks.

Solution Highlights

- AIShield team extracted a pedestrian detection algorithm developed over 10 months and significant investments, within 4 hours & only 4% delta to original model accuracy, as part of an ethical hacking case
- As part of threat mitigation AIShield team integrated a defense model with the original model which brought down the stolen model accuracy to only 15%

Result

Demonstrated the prevention of loss intellectual property & revenue for a computer vision company with impact value of **USD 10 million.**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield
Powered by Bosch

Use Case | Spam Detector and AI Chatbot

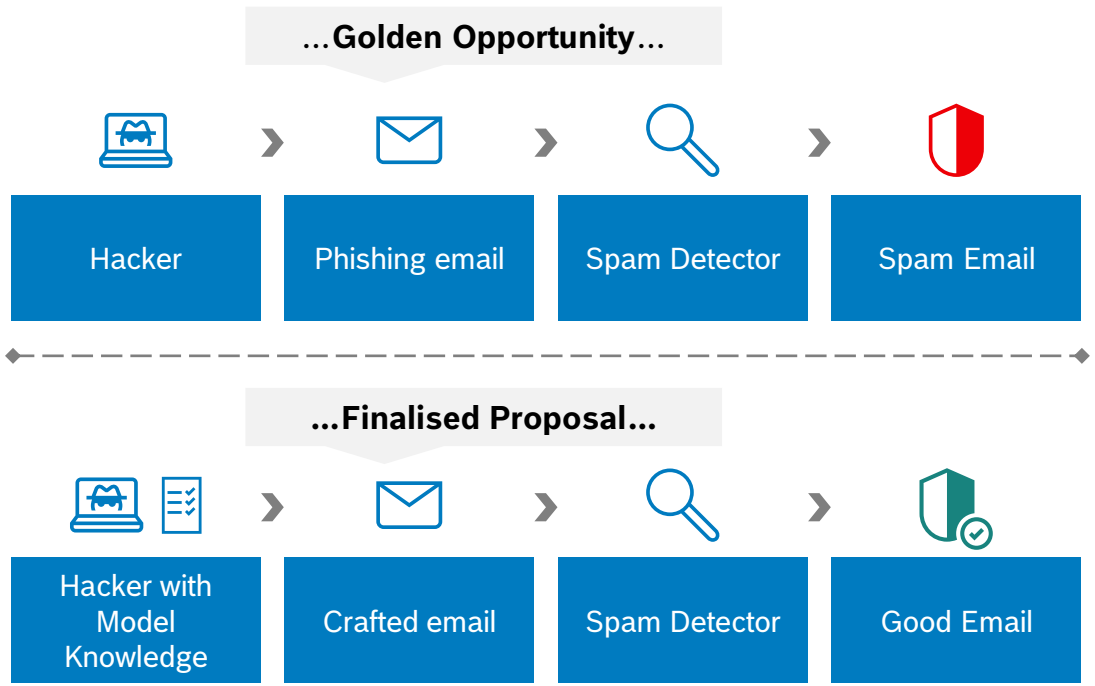
AI Risk Assessment & Mitigation for a Banking company

Algorithm 1 – Spam Email Detection
Algorithm 2 – Conversational AI application

Industry – BFSI (Banking, Financial Services & Insurance)

Customer – American multinational investment bank and financial services holding company

AI Threat Prevented – Model Evasion & Data Poisoning



Background & Objective

The banking organization company desired an overall risk assessment of its ML-powered applications running in various cloud environments including particular interest in ML model evasion & data poisoning.

Solution Highlights

- Assessed the overall risk associated with each ML-powered application using AIShield's AI risk Assessment framework and methodology.
- Iteratively queried ML models running on the cloud through (AIShield's proprietary attack vectors to extract the ML models)
- Evaded email protection system by first building a copy-cat email protection ML model, and using the insights to evade the live system
- Exploited the feedback loop of conversational AI. By repeated interactions using racist and offensive language, biased the dataset towards that language to generate reprehensible material

Result

Prevented Loss of AI Integrity & Brand Reputation

- Provided a risk score for each model for spam detector & conversational AI
- Recommend set of security controls and defense mechanism integration

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield
Powered by Bosch

Use Case | Predictive maintenance in plants

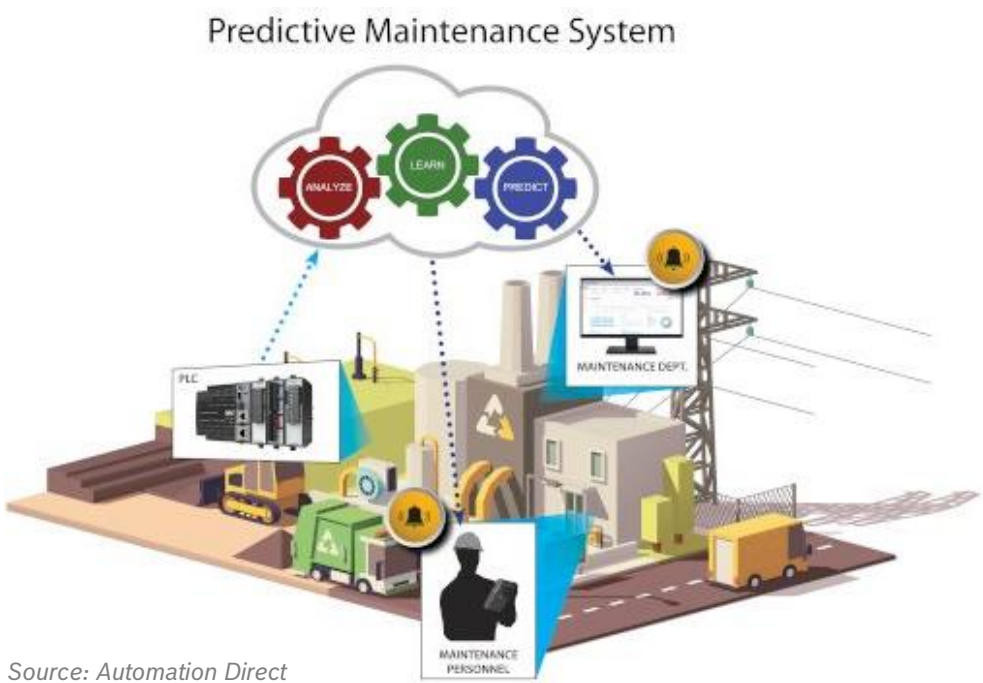
AIoT Risk Assessment for Cloud based IIoT Technologies

Algorithm – Predictive Maintenance

Industry – Manufacturing / Industry 4.0

Customer – APAC based manufacturer

AI Threat Prevented – Model Extraction & Data Poisoning



Source: Automation Direct

Background & Objective

- The manufacturer was concerned about new IIoT technologies introducing new types of cyber threats onto the plant floor, including threats to cloud hosted AI applications
- Manufacturer believed that IP theft has been a primary motive for the cyber-attacks on their factories. State-sponsored cyber-attacks had also reached a critical mass
- Still in the planning & development stage, plant owners wanted to assess risks to their proprietary predictive maintenance solutions

Solution Highlights

- Assessed the overall risk associated with each ML-powered application using AIShield's AI risk Assessment framework and methodology.
- Iteratively queried ML models running on the cloud through AIShield's proprietary attack vectors to extract the ML models
- Demonstrated how the Predictive Maintenance solution could be extracted from the cloud within hours.
- Suggested a defense mechanism for predictive AI application as part of holistic cybersecurity strategy

Result

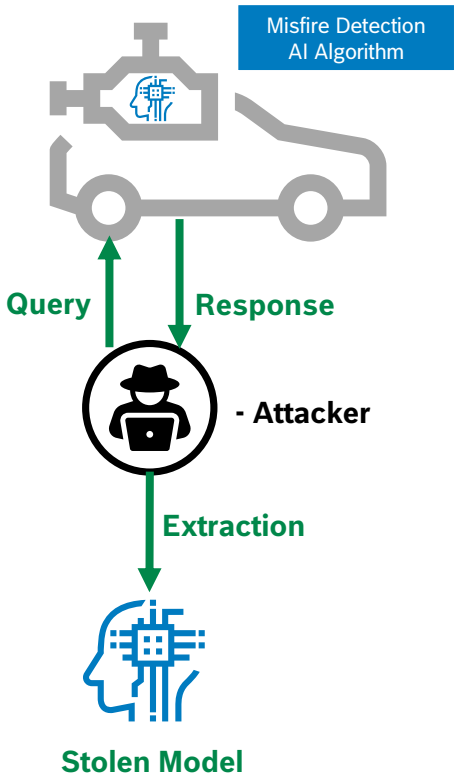
Prevented AI/ML model and associated data related breach of **USD 3 million.**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield Powered by Bosch

Use Case | Misfire Detection

AI Security for IP Protection



Model Extraction Attack for Misfire Detection

Objective

Security hardening of Misfire detection AI Model against stealing attacks

Value Proposition with AIShield

- Prevent loss of **IP and loss of function**

Scope

- Identification & Demonstration of **Model Extraction Attacks**
- Perform **Threat & Risk Analysis** for AI model and system
- Develop **Defense Model** engineered for AI model and Integrate in the target environment

AIShield Outcomes

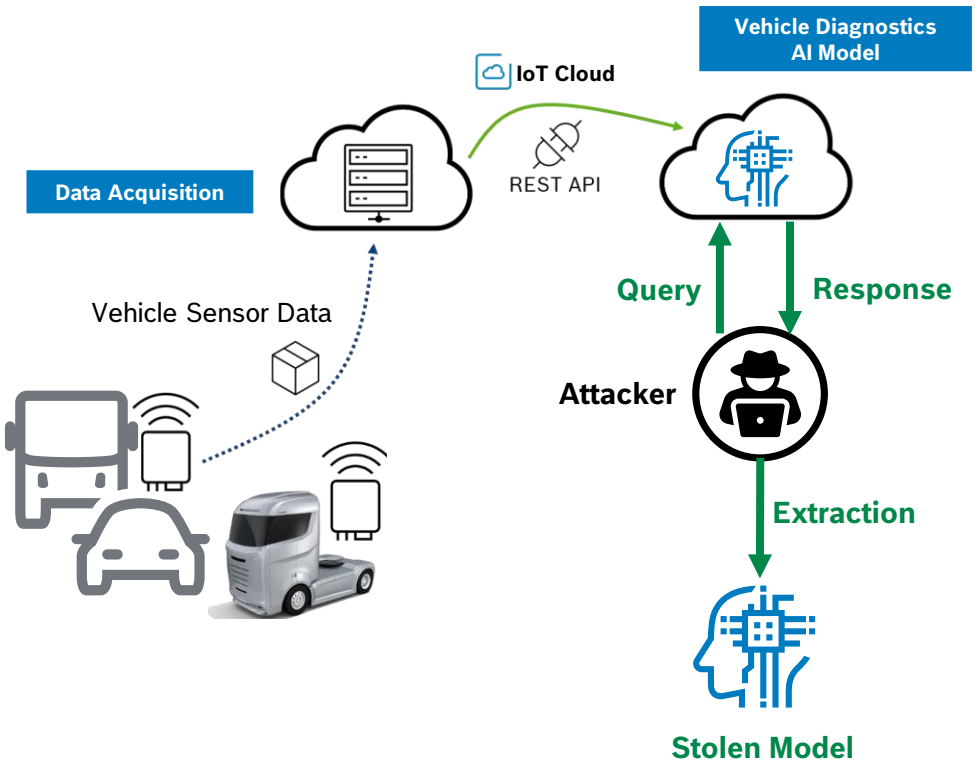
Outcome 1	Demonstration of **Model Extraction Attacks** Vulnerability Report for Model Extraction Attacks (BlackBox & GreyBox)
Outcome 2	**Threat & Risk** Analysis Report for AI Model & System
Outcome 3	Demonstration of **1 Defense Mechanism for Model Extraction Attacks,** **Engineer 1 defense** for Target Embedded application and Demonstrate **Integration** of defense mechanism in **target ECU**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield
Powered by Bosch

Use Case | Vehicle Diagnostics

AI Security for IP protection & prevention of loss of function



How an attacker steals Vehicle Diagnostics AI model

Objective

Security hardening of Vehicle Diagnostics AI Model against stealing attacks

Value Proposition with AIShield

- Prevent loss of **IP and loss of function**

Scope

- Identification & Demonstration of **Model Extraction Attacks**
- Perform **Threat & Risk Analysis** for AI model and system
- Develop threat-informed **defense model** and integrate in the target environment

AIShield Outcomes

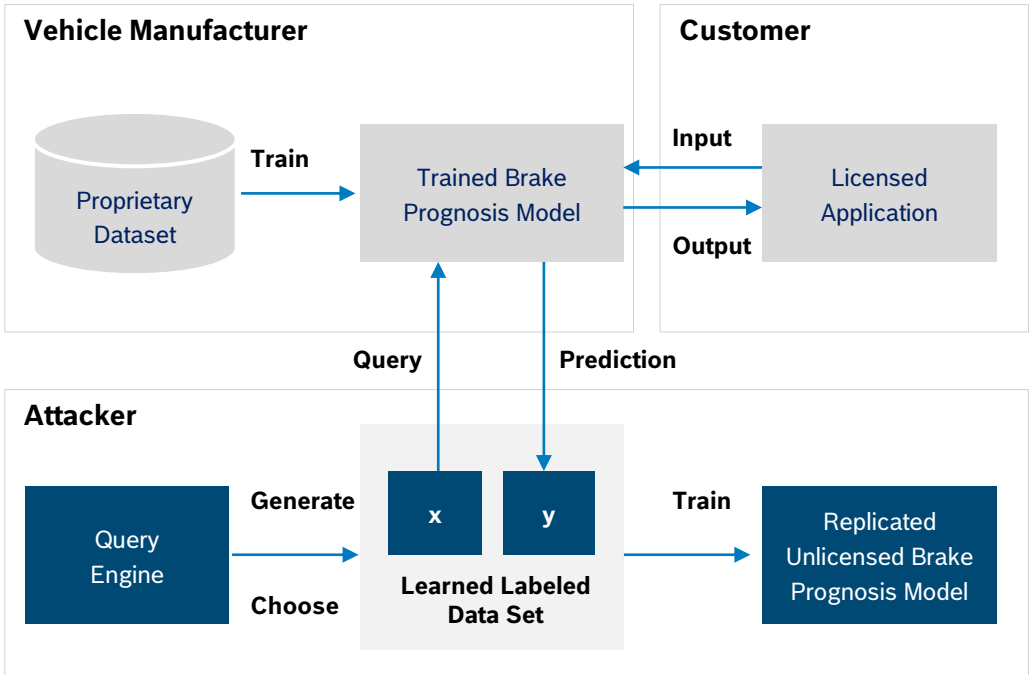
Outcome 1	Demonstration of **2 different relevant attacks** Vulnerability Report for Model Extraction Attacks (BlackBox & GreyBox)
Outcome 2	**Threat & Risk** Analysis Report for AI Model & System
Outcome 3	Demonstration of **1 Defense Mechanism for Model Extraction Attack,** Demonstrate **Integration** of defense mechanism in **Cloud**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield Powered by Bosch

Use Case | Brake Prognosis

Preventing Unlicensed Use of Proprietary AI/ML Model



AIShield team provided AI security threat vulnerability assessment & mitigation for the automotive vehicle manufacturer. The break prognosis model was found vulnerable to model extraction attacks.

Objective

Preventing unlicensed use of proprietary AI/ML model

Value Proposition with AIShield

- Prevent loss of **IP and loss of function**

Scope

- Identification & Demonstration of **Model Extraction Attacks**
- Perform **Threat & Risk Analysis** for AI model and system
- Develop threat-informed **defense model** and integrate in the target environment

AIShield Outcomes

Outcome 1	Demonstration of **2 different relevant attacks** Vulnerability Report for Model Extraction Attacks (BlackBox & GreyBox)
Outcome 2	**Threat & Risk** Analysis Report for AI Model & System
Outcome 3	Demonstration of **1 Defense Mechanism for Model Extraction Attack,** Demonstrate **Integration** of defense mechanism in **Cloud**

© Robert Bosch GmbH 2023. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AIShield Powered by Bosch

Thank You

For more information, please visit

www.boschaishield.com
https://boschaishield.co/guardian

AIShield.contact@bosch.com

AIShield Webpage
AIShield.GuArdIan Webpage

AIShield Intro Video
AIShield.GuArdIan Video

AIShield Brochure

AIShield LinkedIn Page

AIShield on AWS

AIShield - AWS SageMaker Ready

AIShield wins at CES

AIShield recognized at IoTSWC

AIShield at RSAC

AIShield at ET CIO

AIShield at AIIA

AIShield at AIIA MLOps Summit

AIShield at CDMG

AIShield at MedFit

AIShield
Powered by Bosch