



HIDDENLAYER

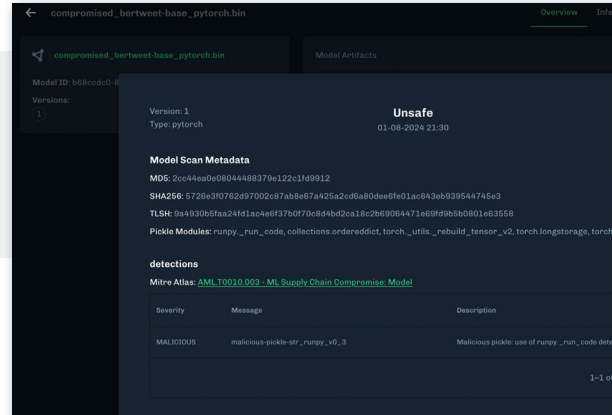
MODEL SCANNER

Third-party and open-source model-sharing repositories have been born out of inherent data science complexity, practitioner shortage & the limitless potential & value they provide to organizations – dramatically reducing the time & effort required for AI adoption. However, such repositories often lack comprehensive security controls, which ultimately passes the risk on to the end user – & attackers are counting on it. The scarcity of security around AI models, coupled with the increasingly sensitive data that AI models are exposed to, means that model hijacking attacks evade traditional security solutions & have a high propensity for damage.

HiddenLayer Model Scanner analyzes models to identify hidden cybersecurity risks & threats such as malware, vulnerabilities & integrity issues. Its advanced scanning engine is built to analyze your artificial intelligence models, meticulously inspecting each layer & components to detect possible signs of malicious activity, including malware, tampering & backdoors.

HiddenLayer Model Scanner is easy to use by simply uploading your model to the Web-based Product Interface or HiddenLayer APIs will automatically analyze it for any security risks. It provides detailed reports on the findings, including recommendations on how to fix any issues & improve the model's security posture.

With HiddenLayer Model Scanner, you can ensure the integrity & safety of your artificial intelligence models, protecting them from any potential cyber threats. Whether you're a data scientist, artificial intelligence engineer, or a business leader, Model Scanner is the essential tool for securing your artificial intelligence assets.



KEY PRODUCT CAPABILITIES

- **Malware Analysis** — Scans AI Models for embedded malicious code that could serve as an infection vector & launchpad for malware
- **Vulnerability Assessment** — Scans for known CVEs & zero-day vulnerabilities targeting AI Models
- **Model Integrity** — Analysis of AI Model's layers, components & tensors to detect tampering or corruption.
- Uses a combination of **supervised learning, unsupervised learning, dynamic/behavioral analysis & static analysis** to deliver detection for a library of adversarial machine learning attacks
- Catalog a **Known-Good State** of your AI Models as a baseline for identifying future tampering
- **Supports a variety of AI Model file types:**
 - Pickle
 - Dill
 - Joblib
 - Numpy
 - Zip
 - ONNX
 - HDF5
 - PyTorch
 - TensorFlow
 - Keras
 - Scikit-Learn
 - Safetensor
 - Cloudpickle

KEY BENEFITS

- Ensure third-party & open source AI models hosted by online communities & repositories are safe & secure to use
- Prevent inheritance of cybersecurity vulnerabilities, malware & corruption via transfer learning of open-source AI Models
- Ensure AI Models are free of vulnerabilities & malware before deploying to production
- Improve the security & integrity of proprietary models & protect your company's intellectual property
- Prevent AI Models from being a launch pad for malware

AVAILABLE ON



Azure Marketplace



AWS Marketplace



Google Cloud Platform

Why HiddenLayer®?

HiddenLayer®, a Gartner recognized Cool Vendor for AI Security, creates security solutions that prevent the latest wave of cybersecurity threats against artificial intelligence assets. Using a patented approach, only HiddenLayer® offers turnkey AI security without requiring increased model complexity, access to sensitive training data, or visibility into the AI assets.

Gartner

COOL
VENDOR
2024