Damage Potential

Table 1: Confidentiality

Metric Value	Description
High	There is a complete breach of confidentiality, leading to the exposure of all resources within the affected AI/ML system to the attacker. Alternatively, the attacker may gain access to certain restricted information, but the revealed data has a direct and significant impact. For instance, the attacker could acquire the administrator's password or the private encryption keys of a web server, resulting in serious consequences.
Low	There is a partial compromise of confidentiality. The attacker manages to gain access to certain restricted information; however, they lack control over the specific data obtained, and the extent or type of loss is restricted. The disclosed information does not result in a direct and significant loss to the affected AI/ML system.
None	There is no loss of confidentiality within the impacted AI/ML system.

Table 2: Integrity

Metric Value	Description
High	There is a complete compromise of integrity or a total breakdown of protection within the AI/ML system. For instance, the attacker gains the ability to modify any or all files safeguarded by the system. Alternatively, although only certain files can be modified, any malicious alterations would result in direct and significant consequences for the affected system.
Low	It is possible to modify the data within the AI/ML system; however, the attacker lacks control over the consequences of the modifications, or the extent of modification is restricted. The data alterations do not result in a direct and significant impact on the affected system.
None	There is no loss of integrity within the impacted AI/ML system.

Table 3: Availability

Metric Value	Description
High	There is a complete loss of availability within the AI/ML system, enabling the attacker to fully deny access to its resources. This loss of availability can either be sustained, meaning the attacker continues to deliver the attack, or persistent, where the condition persists even after the attack has concluded. Alternatively, the attacker possesses the capability to partially disrupt availability, but this loss of availability has a direct and serious impact on the affected system. For example, the attacker may be unable to disrupt existing connections but can prevent new connections, or they can repeatedly exploit a vulnerability that, although each successful attack only leaks a small amount of memory, eventually leads to a complete unavailability of the service.
Low	The performance of the AI/ML system is diminished, or there are intermittent disruptions in the availability of its resources. Even if the vulnerability can be repeatedly exploited, the attacker lacks the capability to entirely block access for legitimate users. The resources within the affected system are either partially available consistently or fully available intermittently. However, overall, there is no immediate and significant impact directly affecting the affected system.
None	There is no impact to availability within the impacted AI/ML system.

Exploitability

Table 1: Attack Vector

Metric Value	Description
Network	The AI/ML system is susceptible to remote attacks originating from the internet. This type of vulnerability, referred to as 'remotely exploitable,' enables targeting the system from a distance, potentially spanning multiple routers. An example of a network attack is an attacker causing a denial of service (DoS) by sending a specially crafted TCP packet across a wide area network (e.g., CVE-2004-0230).
Adjacent	The AI/ML system is network-connected, but the attack is constrained to a specific level within the network structure. This limitation implies that the attack can only occur from a logically adjacent area, which may include the same physical or logical network, such as Bluetooth or a local IP subnet. Additionally, the attack could originate from within a secure or restricted administrative domain, such as MPLS or a secure VPN connecting to an administrative network zone. One example of an Adjacent attack would be an ARP (IPv4) or neighbor discovery (IPv6) flood leading to a denial of service on the local LAN segment (e.g., CVE-2013-6014).
Local	The AI/ML system is not directly connected to the network, and the attacker can exploit the vulnerability either locally (e.g., keyboard, console), or remotely (e.g., SSH).
Physical	The attack in this case necessitates the physical interaction with the AI/ML system which can be a brief interaction (e.g., an evil maid attack). An example of such an attack is a cold boot attack in which an attacker gains access to disk encryption keys after physically accessing the target system. Other examples include peripheral attacks via FireWire/USB Direct Memory Access (DMA).

Table 2: Attack Complexity

Metric Value	Description
Low	There are no special access conditions or exceptional circumstances required for the attack. An attacker can expect a successful attack when targeting the AI/ML system.
High	For a successful attack to occur, certain conditions must be met that are beyond the control of the attacker. This means that the attacker cannot simply launch the attack whenever they want. Instead, they must invest a significant amount of effort in preparation or execution against the AI/ML system before they can expect the attack to be successful. For example, a successful attack may depend on an attacker overcoming any of the following conditions: • The attacker must gather knowledge about the environment in which the vulnerable target/component exists. For example, a requirement to collect details on target configuration settings, sequence numbers, or shared secrets. • The attacker must prepare the target environment to improve exploit reliability. For example, repeated exploitation to win a race condition, or overcoming advanced exploit mitigation techniques. • The attacker must inject themselves into the logical network path between the target and the resource requested by the victim in order to read and/or modify network communications (e.g., a man in the middle attack).

Table 3: Privileges Required

Metric Value	Description
None	The attacker, being unauthorized prior to the attack, does not necessitate any access to settings or files of the AI/ML system in order to execute an attack.
Low	The attacker requires privileges that grant basic user capabilities, typically limited to modifying settings and files owned by a user. Alternatively, an attacker with low privileges has the ability to access only non-sensitive resources.
High	The attacker requires privileges that grant substantial control (e.g., administrative privileges) over the AI/ML system, enabling access to its settings and files.

Table 4: User Interaction

Metric Value	Description
None	The vulnerable AI/ML system can be exploited without requiring any user interaction.
Required	For successful exploitation of the vulnerability, a user needs to perform a specific action before the vulnerability can be exploited. For instance, a successful exploit may only occur during the installation of an application by a system administrator.

Table 5: Scope

Metric Value	Description
Unchanged	An exploited vulnerability can exclusively impact resources that fall under the management of the same security authority. In this scenario, the vulnerable AI/ML system and the affected components are either identical or both are under the management of the same security authority.
Changed	An exploited vulnerability can have an impact on resources that extend beyond the security scope managed by the security authority of the vulnerable AI/ML system. In this scenario, the vulnerable system and the affected components are distinct and managed by different security authorities.

Table 6: Exploit Code Maturity

Metric Value	Description
Not Defined	Assigning this value indicates there is insufficient information to choose one of the other values, and has no impact on the overall Temporal Score, i.e., it has the same effect on scoring as assigning High.
High	There is functional autonomous code available, or in some cases, no exploit is required as the vulnerability can be triggered manually. Comprehensive details regarding the vulnerability are widely accessible. The exploit code is effective in every situation, or it is actively being disseminated through an autonomous agent such as a worm or virus.
Functional	There is available functional exploit code that effectively operates in various scenarios where the vulnerability is present. The code demonstrates its efficacy across a wide range of situations where the vulnerability is exploitable.
Proof-of-Concept	Proof-of-concept exploit code is accessible, or in some cases, an attack demonstration may not be feasible for most systems. However, it is important to note that the provided code or technique may not be fully functional in all situations and could necessitate significant modifications by a skilled attacker.
Unproven	No exploit code is available, or an exploit is theoretical.

Affected Users

Table 1: Affected Users

Metric Value	Description
All users	The vulnerability in the AI/ML system affects all users which can lead to significant consequences for everyone involved.
Some users	The vulnerability in the AI/ML system has an impact on select users, targeting specific subsets within the user base and resulting in limited consequences for that particular portion.
One or small group of users	The vulnerability in the AI/ML system is targeted or localized, typically affecting a single individual or a small number of people.

Discoverability

Table 1: Discoverability

Metric Value	Description
High	It is unlikely to easily find the vulnerability in the AI/ML system and it requires a remarkable amount of time and effort.
Medium	Detecting the potential for malicious exploitation in this vulnerability in the AI/ML system requires thoughtful consideration and analysis, as it operates at a level that may not be immediately apparent.
Low	The vulnerability in the AI/ML system can be easily identified through a few basic observations, posing minimal challenge in its detection.

Ease of Mitigation

Table 1 Responsible Entity

Metric Value	Description
Security Researcher	The main responsibility of the security researchers lies in responsibly disclosing vulnerabilities and ensuring the information reaches the appropriate entity. The risk in this process can be the potential miscommunication, delay in reporting, or incomplete understanding of the vulnerability in the AI/ML system in comparison to other responsible entities, which could hinder the mitigation process.
Third-party Vendor	If the vulnerability within the AI/ML system can be exploited using any of its components, it becomes the responsibility of the third-party vendor, acting as the manufacturer, to either enhance the design of the component or replace it with a more reliable alternative.
Technology Vendor	Failure to promptly release patches or updates to fix vulnerabilities can leave users exposed to potential attacks. With the complete knowledge about the design of the AI/ML system, vendors must thoroughly investigate

	reported vulnerabilities, develop effective patches to ensure proper mitigation.
System Administrator	If administrators – as the frontline entity to deal with the vulnerabilities - fail to promptly apply patches or implement necessary security measures, the risk of successful attacks targeting the vulnerabilities increases. Therefore, their actions directly impact the mitigation of the vulnerabilities in the AI/ML system they manage.

Table 2: Remediation Level

Metric Value	Description
Not Defined (X)	Assigning this value indicates there is insufficient information to choose one of the other values, and has no impact on the overall Temporal Score, i.e., it has the same effect on scoring as assigning Unavailable.
Unavailable (U)	Either a feasible solution does not exist or, if one does, it proves impossible to implement.
Workaround (W)	An unofficial, non-vendor solution exists, where users or administrators of AI/ML may take matters into their own hands by creating a patch or providing steps to mitigate the vulnerability. In certain cases, individuals within the user community offer alternative measures or workarounds to address or minimize the impact of the vulnerability.
Temporary Fix (T)	An official but temporary solution is available, which may include the vendor providing a temporary hotfix, tool, or workaround. In such cases, the vendor acknowledges the issue and offers a temporary measure to address the vulnerability until a permanent fix can be developed and implemented.
Official Fix (O)	A comprehensive vendor solution is available, either in the form of an official patch released by the vendor or through an available upgrade. This means that the vendor has provided a complete resolution to address the vulnerability, offering users or administrators the means to apply an official fix or update to mitigate the risk effectively.