A HYBRID MACHINE LEARNING APPROACH FOR DETECTING INTENTIONAL AND UNINTENTIONAL INSIDER THREATS WITH MITIGATION THROUGH BEHAVIORAL BIOMETRICS AND USER PROFILING MECHANISM

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1. INTRODUCTION TO INSIDER THREAT

In an organization, many security threats are prevalent, and one of the challenging security threats is the Insider threat. The Insider threat is one of passive security threats that initiated by an employee who exhibits trusted authorization within organization to critical resources, systems and data by violating internal security policies and terms (Yuan & Wu, 2021). An insider may be an employee, contractor, or business partner who exploits legitimate privileges such as trust, access, security policies, or knowledge to access sensitive information. They perform unauthorized activities such as data theft, sabotage, misuse of resources, data exfiltration and accidental leaks to compromise the organization's security.

In the view of Nurse et al. (2014), an insider is classified into intentional and unintentional insiders based on the nature of abnormality. Intentional insiders are motivated to pose harm to a system, and unintentional insiders perform malicious activities due to their carelessness. However, both types of threats possess major potential consequences to an organization, such as security breaches, data breaches, financial losses and reputation damage. These challenges possess the need to detect and mitigate both intentional and unintentional insider threats. Past research shows less interest in detecting and mitigating both intentional and unintentional insiders. Thus, detection and mitigation of both intentional and unintentional insiders is required.

1.1.1. INSIDER

There are many definitions available for an insider in the literature and reports. Some of the significant definitions are:

- Waiganjo and Nandjenda (2025) described an insider as an individual who intends to pilfer intellectual property, interrupt reliable networks, and use information previously preserved in portable devices to engage in fraudulent scams.

- Trivedi et al. (2025) stated that an insider is a trusted user who abuses their authority, knowledge, and connections within a secure network to prompt significant damage by accessing an internal system of an organization.
- IBM (2024) defined an insider as an aggrieved employee who deliberately exploits confidential access with legitimate access for personal retaliation or fiscal gain to interfere with the company's performance.
- Kamatchi and Uma (2024) defined an insider as an authorized individual endorsing lawful rights based on acquiring, depicting, or managing organizational assets.
- Yuan and Wu (2021) emphasized an insider as an individual who gains unrestricted authority to captivate implicit computing resources, including the computer architecture, information, or software, in an unpredictable manner due to their known acquittances.
- Bishop et al. (2009) characterized an insider as a credible entity with a demarcated set of rules and policies.

Considering the above definitions, it is inferred that an insider can be defined as an individual who utilizes their authority to access an organisation's sensitive information either intentionally or unintentionally.

1.1.2. INSIDER THREAT

At present, recent literature highlights many definitions that are available for an insider threat. Some of significant definitions of insider threat are listed below:

- Palmer (2025) defined an insider threat as the mechanism of performing insider activities that are susceptible to confidential information including resources, data, and adverse the functioning of a corporate organization.
- Wajganjo and Nandjenda (2025) described an insider threat as the authoritative insider who performs vulnerable activities that intensively distress an organization.
- Alzaabi and Mehmood (2024) highlighted that an insider threat is a mechanism of abusing the vulnerability in the security system of an organization.

- Whitelaw et al. (2024) revealed the behavioural indicators of insiders, such as insider personality traits, social indicators, insider characteristics, psychosocial, insider behavioral approaches and motivation, sociotechnical indicators, and psychopathy.

From the above, it is derived that an insider threat can be denoted as a security threat within an organization, and it can be achieved by an individual misusing their legitimate authority while accessing sensitive information.

1.1.3 CLASSIFICATION OF INSIDER THREAT

Insider threat is classified into two categories: Intentional insider threat and Unintentional insider threat based on the intention of attack, and are illustrated in Figure 1.1.

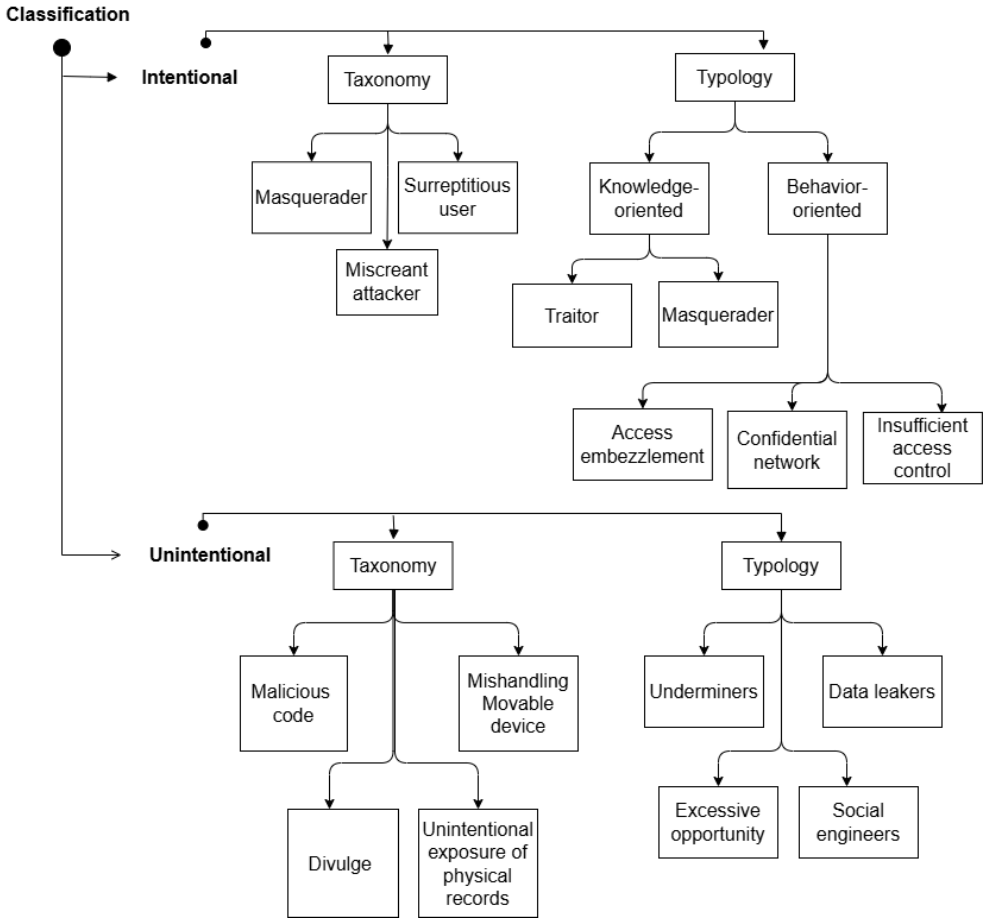

Figure 1.1 Classification of Insider Threat

i) Intentional Insider Threat

Intentional insider threat is instigated by a former or current employee acquiring absolute privileges and performing unauthorized activities by intentionally manipulating secure information (Hunker & Probst, 2011) and endangering an organization's security policy. According to Cole and Ring (2005), intentional insiders are categorized based on two constraints: Action performed and Potential consequences, as shown in Figure 1.2.

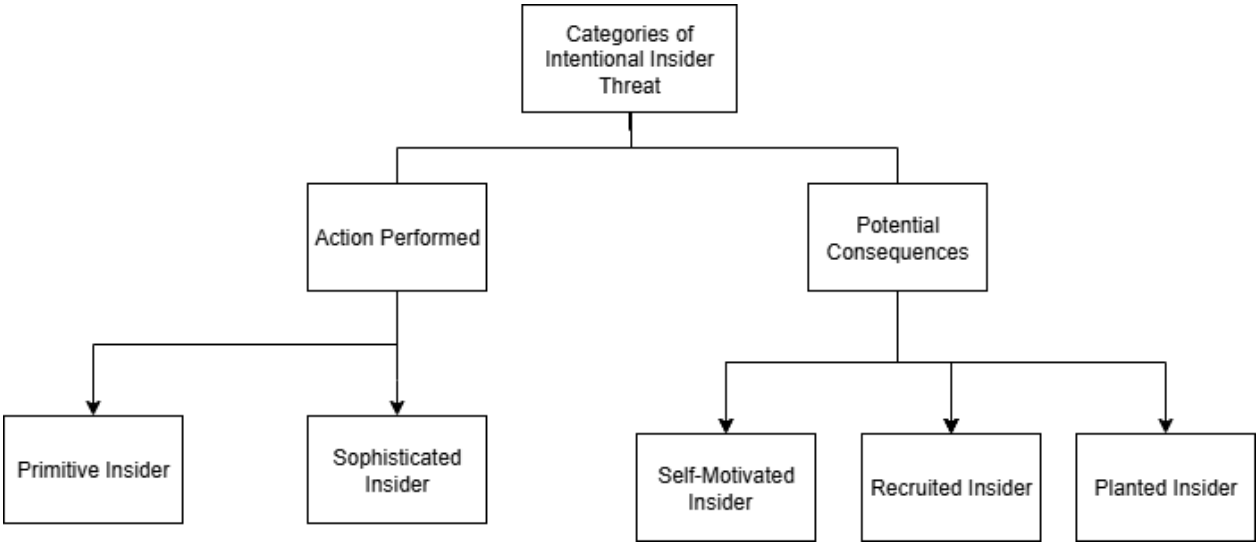

Figure 1.2: Categories of Intentional Insider Threat

Insiders are profiled further as primitive or sophisticated based on actions, existence, and qualities of work acquittances. On the contrary, intentional insider threats are categorized into self-motivated insider, recruited insider, and planted insider based on potential consequences. The basic intention of intentional insider threat is to confiscate a narrow passage of destruction in terms of fiscal, political and personal. The intentional insider threats can also be categorized based on taxonomy and typology. The intentional insiders are classified into masquerader, miscreant attacker, and surreptitious users based on taxonomy. In terms of typology, the intentional insiders are classified into behavior oriented and knowledge oriented. The following table 1.1 analyses the intentional insider based on taxonomy and typology.

Table 1.1: Analysis on Taxonomy and Typology of Intentional Insider Threats

Categorization on	Types	Definition	Key Characteristics	Consequences
Taxonomy	Masquerader	An intentional insider utilizes extrinsic attackers to infiltrate the network or confines sensitive credentials through intrinsic attackers.	To confine the network	Data loss
	Miscreant attacker	An intentional insider who corrupts the trusted authority for embezzlement.	To misgovern the legitimate authorization	Reputation loss
	Surreptitious users	An intentional insider who manages the confidential network utilizing predominant influence.	To misuse the authority of others	Reputation loss
Typology	Behavior oriented	The intentional insider possesses the major behavior characteristics of embezzling the access, evading the network, or insufficient access control (Bellovin, 2008).	Access embezzlement, confidential network evading, and insufficient access control	Reputation loss
	Knowledge oriented	The intentional insider is one with distinct knowledge characteristics such as being a masquerader to possess a targeted user's credentials or a traitor who intends to exploit the sensitive information (Salem et al., 2008).	Being a traitor or a masquerader	Data loss

Table 1.1 describes that the categorization or taxonomy of Intentional Insider is accomplished on the basis of discerning secured networks using unauthorized audit logs in terms of masquerader, miscreant attacker, and surreptitious users. However, the typologies of Intentional insiders are based on characteristics of behavior and knowledge.

ii) Unintentional Insider Threat

An unintentional insider threat is induced by working individual, business affiliates, or contractors who carelessly acquire control over sensitive data, a system, or a network in an organization. It leads to sabotage without malignant intention, as it affects an organization's Confidentiality, Integrity, and Availability (CIA). Khan et al. (2022) acknowledge that unintentional insiders are considered as inadvertent for their peculiar behaviour of instinctively exploiting computational networks through appropriate access. It simplifies the disclosure of confidential information unintentionally (Raskin et al., 2010). Further, Homoliak et al. (2019) groups the unintentional insiders based on Privacy rights and Nature of negligence with their vulnerability to data exposure, which is illustrated in figure 1.3.

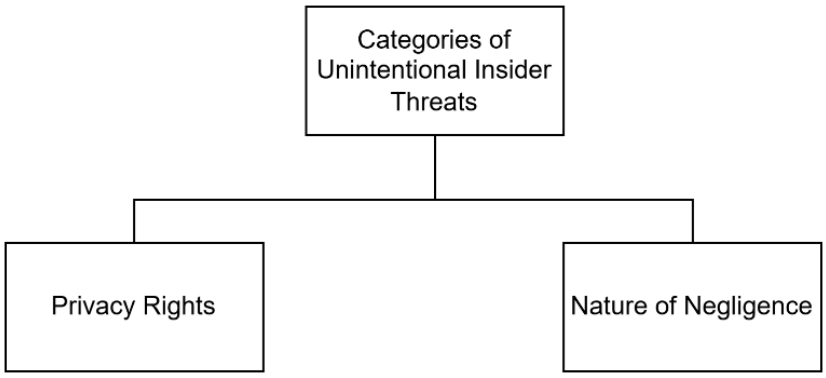

Figure 1.3: Categories of Unintentional Insider Threat

The unintentional insider threats are also categorized based on taxonomy and typology. Based on taxonomy, the unintentional insider threats are classified into malicious-code, divulge, unintentional exposure of physical records, and movable device previously owned. The typology of unintentional insiders is underminer, excessive opportunistic, data leaker and social engineer. The following table 1.2 analyses the unintentional insider based on taxonomy and typology.

Table 1.2: Analysis on Taxonomy and Typology of Unintentional Insider Threats

Categorization on	Types	Definition	Key Characteristics	Consequences
Taxonomy	Malicious-code	An unintentional insider unknowingly evolves the malware through sensitive information obtained via social engineering, including phishing attacks and embedded removable devices.	To develop the malware	Data loss
	Divulge	An unintentional insider facilitates the confidential information either on the internet via the Web or isolated via mail-fax to an unauthorized receiver.	To disseminate sensitive information	Data loss
	Unintentional exposure of physical records	An unintentional insider who divulges the disconnections by neglecting or absquatulating the documents containing sensitive records.	To reveal the information	Data loss
	Movable device previously owned	An unintentional insider acquires the computational records of being scrutinized, exposed, or receded with a CD, laptop, or removable drive.	To gather the information	Data loss

Categorization on	Types	Definition	Key Characteristics	Consequences
Typology	Underminer	The unintentional insider infringes the privacy policies.	To invade the security policy	Reputation loss
	Excessive opportunistic	The unintentional insider tries to trespass security policies and measures to achieve more efficiency.	To intrude on the security network	Reputation loss
	Data leakers	An unintentional insider who exposes sensitive information.	To disclose the data	Data loss
	Social engineers	An unintentional insider who compromises security by exploiting human psychology and trust.	To conceal the security	Reputation loss

Table 1.2 shows that the taxonomy of unintentional insider consists of malicious-code, divulge, unintentional exposure of physical records, and movable devices previously owned. The typologies of unintentional insiders include underminer, excessive opportunistics, data leakers, and social engineers based on carelessness-led data exposure.

1.2. ATTACK INDICATORS

Attack indicators are the contributing factors of insider driven security breaches where illicit authority is acquired for gaining confidential services that are achieved through traitors and masqueraders utilizing their contingent attack indicators (Salem et al., 2008). Some of the well-known attack indicators of insider threats are

- Phishing emails: Hijack the security network using stolen credentials (Cleghorn, 2013; Saxena et al., 2020),
- Privilege escalation: Exploits the vulnerability to bypass the security protocol (Jaafar et al., 2016),

- Data exfiltration: Conduct unlawful data replication to access sensitive information (Janssen, 2013),
- Advanced Persistent Threat (APT): Perform spear-phishing to a targeted employee (Giura & Wang, 2012).

1.3. ANALYSIS OF DIFFERENT TYPES OF INSIDER THREAT

An intentional or unintentional insider threat, each type of insider threat can cause different potential threats. A legitimate employee with authorization is considered legitimate as he/she performs malicious activities. Unintentional insiders are prone to phishing, and social engineering attacks due to carelessness are hard to detect because of their trustworthiness. The different types of insider threats are analyzed in terms of definition, detection, intention, and example in following table 1.3.

Table 1.3 Analysis of Different Types of Insider Threats

S.no	Category	Intentional insider	Unintentional insider
1.	Definition	A current or former employee who deliberately exploits their authorization to access the security system of organization for malicious purposes.	A negligent employee who inadvertently compromises the security of an organization.
2.	Detection	Hard to detect because of intentional concealment and utilization of legitimate access.	Due to the human error and carelessness, it is challenging to obstruct without thorough training and awareness.
3.	Intention	Deliberately cause harm to the organization in terms of monetary and reputation.	Not intended to perform malicious activities.
4.	Examples	Counterfeiting of intellectual property, IT sabotage, deception, or data loss.	Phishing attacks, unauthorized disclosure of confidential information by error, or losing a device containing confidential information.

Both types of insider threats cause potential threats to an organization. So, it is essential to detect and mitigate the insider threat.

1.4. POTENTIAL CONSEQUENCES OF INSIDERS

The potential challenges confronted by both intentional and unintentional insiders can substantially affect the security measures in an organization (Saxena et al., 2020). Intentional insiders specifically presume malignant activities by individuals involve legitimate authority, intended to sabotage the organization, burgle confidential information, or interrupt the processes. On the contrary, unintentional insiders often emanate from negligence, for being oblivious, or perform inadvertent activities that conciliate security via phishing links or misusing confidential information.

Figure 1.4 illustrates a comprehensive illustration of the potential consequences of both intentional and unintentional insiders, emphasizing various scenarios in which insider threats could manifest. The implications of insiders and the comprehensive requirement for addressing them based on an organization's security strategies are briefly discussed in this section.

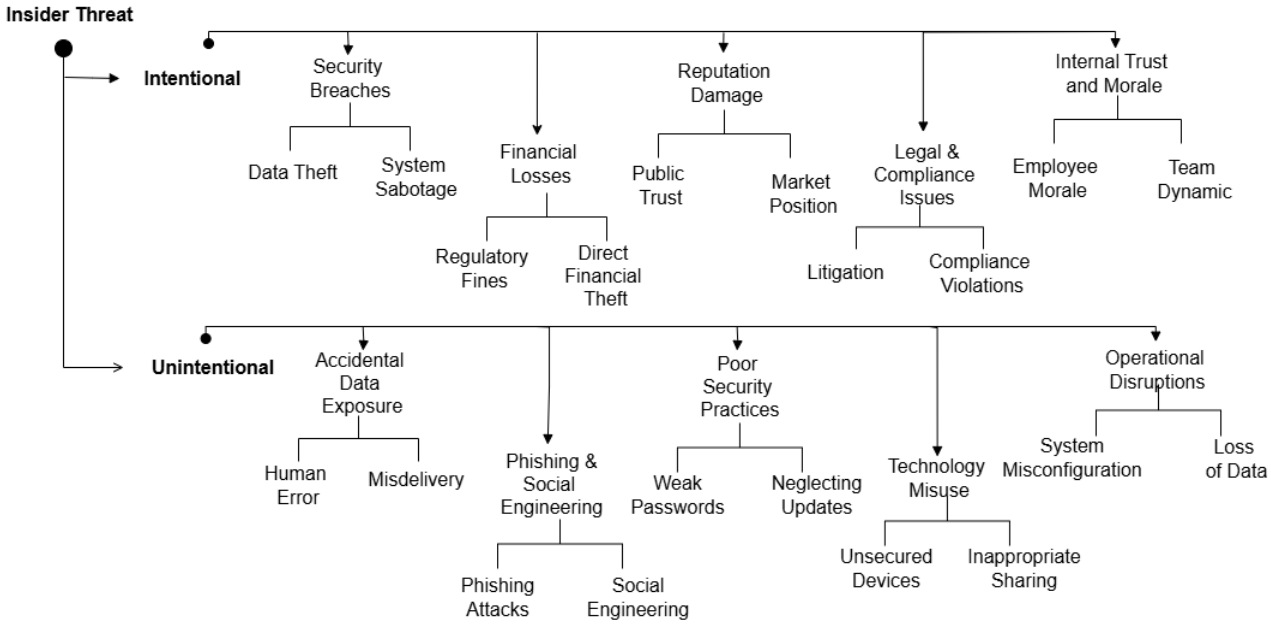

Figure 1.4: Categorization of the Potential Consequence of Intentional and Unintentional Insiders

1.4.1 POTENTIAL CONSEQUENCES OF INTENTIONAL INSIDERS

Intentional insiders cause a potential cybersecurity threat as they exploit legitimate access to compromise systems, steal sensitive data, or disrupt operations. Their actions can lead to severe explicit and implicit threats that result in potential consequences of intentional insiders. The potential consequence of intentional insiders is higher in terms of security, operations and trust in an organization. An analysis of significant consequences of intentional insiders, types, definition, and notable indicators are discussed in table 1.4.

Table 1.4 Significant Consequences of Intentional Insider Threats

Consequences	Types	Definition	Notable Indicators
Security Breaches	Data theft	Intentional insiders may steal information, and embezzle trade secrets result in erosion of organization's competitive edge.	Steal proprietary and customer information, and embezzling trade secrets
	System Sabotage	Insiders who deliberately cause system damage and malware introduction into the system, could cause the obliteration of key data.	Cautious system damage
Financial Losses	Regulatory Fines	An insider-driven data breach could result in regulatory fines.	Fines
	Direct financial theft	An insider may embezzle through unauthorized transactions led to immediate financial losses and required complex reconciliation efforts.	Unauthorized transactions
Reputation damage	Public trust	Insider driven data breach can severely damage an organization's reputation among public and stakeholder confidence.	Reputation loss
	Market	An insider driven data breach	Damaging stock

Consequences	Types	Definition	Notable Indicators
	position	causing the adverse media coverage could affect stock prices and market perception, substantially leading to brand damage.	price and reputation
Legal and Compliance Issues	Litigation	Insider driven breach would cause customers to file lawsuits or regulatory bodies may impose legal proceedings against the organization.	File lawsuit
	Compliance Violations	Insider driven breach led to compliance violence due to bypassing security controls.	Violating the compliance
Internal trust and morale	Employee morale	The insider threat detection in an organization can lower employee morale	Decreased moral among employee
	Team dynamic	Insider threat can disrupt team dynamics when suspicion remains persistent.	Affects the team performance

Table 1.4 shows that intentional insider threats result in major security breach with vast data loss in an organization.

1.4.2 POTENTIAL CONSEQUENCES OF UNINTENTIONAL INSIDERS

An unintentional insider who carelessly exposes sensitive information would result in major consequences such as data breaches, fiscal loss, and brand damage for an organization. An analysis of significant consequences of unintentional insiders, types, definition, and notable indicators are briefed in table 1.5.

Table 1.5 Significant Consequences of Unintentional Insider Threats

Consequences	Types	Definition	Notable Indicators
Accidental Data Exposure	Human Error	Unintentional insider may accidentally share sensitive information via unsecured channels such as phishing attacks and social engineering.	Accidental data exposure
	Misdelivery	Unintentional insiders are prone to weak passwords and lack of multi-factor authentication that could increase the risk of unauthorized access.	Lack of consciousness
Technology Misuse	Unsecured Devices	Unintentional insider utilizing unsecured devices such as personal device without essential security measures and lack of encryption lead to data theft.	Lack of security and advanced encryption techniques
	Inappropriate sharing	Unintentional insider shares the sensitive information through unsecured platforms.	Unauthorized data sharing
Operational Disruptions	System misconfiguration	Unintentional insiders are prone to software applications with misconfiguring systems and insufficient security controls which can create vulnerabilities or disrupt operations.	Vulnerable software and security systems
	Loss of data	Unintentional insiders delete critical data without backup could disrupt business processes and require costly recovery efforts.	Damaging sensitive information

Table 1.5 highlights that unintentional insider threats is majorly caused by negligence and result in sensitive information loss in an organization. Both intentional and unintentional insider threats cause potential challenges to organizations, necessitating robust security measures, continuous monitoring, and a culture of security awareness. Addressing these challenges requires a multifaceted approach to detect and mitigate risks which in turn will protect the organizational assets.

1.5. PROBLEM JUSTIFICATION

Detection and mitigation of both intentional and unintentional insider threats are very important. Recent statistics that are tabulated in table 1.6 discusses the potential consequences of insider threats which justifies the selected research problem.

Table 1.6: Recent Statistics on the Consequence of Insider Threat

Source	Recent statistics on consequences of insider threat	Threat caused	Consequences
Forbes 2024	Insiders accompanied by Lapsus (hacker group) to achieve organizational malice, spying, and proliferation of business targets by 2023 (Sayegh, 2023).	organizational malice, spying, and prognosticate	Increased business targets.
2024 Annual Data Exposure Report by Code42	28% increase in insider-driven data loss since 2021 (Mimecast, 2024)	Advertent malignant activities	Increased information loss
2024 Cost of Insider Threat Global Report by Ponemon Institute	$15.38 million cost of insider threat incidents where unintentional insiders are responsible for 56% of incidents (Proofpoint, Inc., 2022)	Advertent malignant activities and carelessness	Monetary loss
IBM 2024 Cost of a Data Breach Report	An insider-caused data breach costs $11.45 million, taking over 250 days to identify and 80 days to mitigate them (IBM, 2024).	Advertent malignant activities	Data breach

Source	Recent statistics on consequences of insider threat	Threat caused	Consequences
2024 Verizon Data Breach Investigation Report	35% of increased breaches, with 73% from miscellaneous errors. 68% of unintentional driven data leaks (Verizon, 2024).	Information misconfiguration, misdelivery	Data breach
Verizon Data Breach Investigation Report 2023	96% of insider activities for personal gain (Verizon, 2023)	Advertent malignant activities	Data breach

The above table collectively emphasises the urgent need for organizations to strengthen their security defences, particularly against insider threats, through effective detection and mitigation strategies. Hence, an attempt has been made to develop an approach that detects and mitigates intentional and unintentional insider threats.

1.6. APPLICATION AREA OF INSIDER THREAT DETECTION AND MITIGATION

The application areas of insider threat detection and mitigation are significant across diverse fields and are intended to secure confidential data, preserve functional integrity, and assure adherence to administrative policies (Al-Mhiqani et al., 2024). The application areas are:

1. Transportation
2. Healthcare
3. Nuclear Facilities
4. Gas and Oil
5. Energy
6. Water
7. Smart city
8. Industrial Automation
9. Cloud and
10. Wherever user logs are prevalent

The above-mentioned application areas emphasize the requirement of employing an intelligent insider threat detection and mitigation approach to protect secured data by ensuring the safety and prolongation of various sector functions.

1.7. PROBLEM STATEMENT

To develop an approach that detects and mitigates both intentional and unintentional insider threats.

1.8. OBJECTIVES OF THE THESIS

The primary objective of the thesis is to detect and mitigate both intentional and unintentional insider threats.

The secondary objectives formulated in the thesis are as follows.

- To detect and classify both intentional and unintentional insider threats by effectively addressing the challenges of class imbalance and reducing the misclassification rate.
- To mitigate unintentional insiders by enhancing detection accuracy and minimize the error rates in the user authentication system.
- To mitigate intentional insiders through user profiling mechanism techniques.

1.9. SIGNIFICANT CONTRIBUTIONS OF THE THESIS

The significant contribution of the thesis work is illustrated in Figure 1.5:

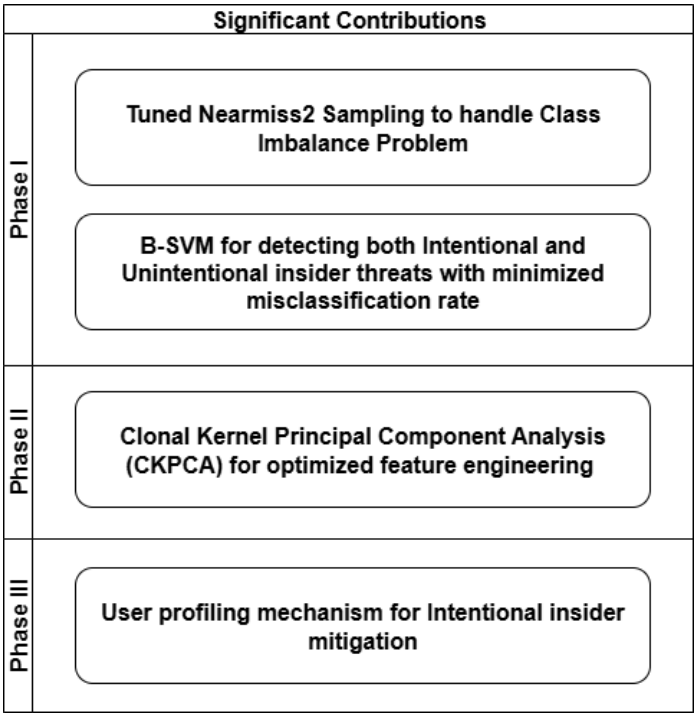

Figure 1.5: Significant Contributions

i) *Phase I-*

- *Enhanced Nearmiss-2 for sampling*: Implemented and tuned the sampling technique for significantly handling the class imbalance problem.
- *B-SVM for classification*: The development of a B-SVM algorithm with an Support Vector Machine (SVM) for initial detection of intentional insiders and a Balanced Iterative Reducing Clustering Hierarchy (BIRCH) algorithm for refining classifications by re-evaluating false positives, resulting in enhanced detection accuracy and minimal misclassification rates.

ii) *Phase II*

- *CKPCA for feature engineering*: Integrated Clonal selection algorithms, Kernel mean embedding, and Principal Component Analysis (CKPCA). Applied CKPCA with Deep Belief Network to build a user classification model for behavioral biometrics for improved detection accuracy and reduced EER.

iii) *Phase III*

- *Insider Profiling*: The intentional insiders are profiled to achieve overall system security by continuously updating user statuses against insiders, which provides a significant contribution to the field of insider threats.

1.10. ORGANIZATION OF THE THESIS

The thesis is structured in the following order,

- Chapter 2 deals with existing works done by the researchers in the field of insider threat detection and mitigation and the research gaps identified.
- Chapter 3 provides an outline of the proposed framework for insider threat detection and mitigation.
- Chapter 4 describes the phase I of proposed methodology namely Preprocessing and Insider Detection (P&ID) with experimental results.
- Chapter 5 elaborates the phase II of proposed methodology namely Unintentional Insider Mitigation (UIM) with experimental results.
- Chapter 6 discusses the phase III of proposed methodology namely Intentional Insider Mitigation (IIM) with experimental results.

- Chapter 7 summarizes and concludes the entire research work.
- Chapter 8 provides the future recommendations.

1.11. CHAPTER SUMMARY

Chapter 1 overviewed insider threats and their taxonomies, the possible consequences of intentional and unintentional insider threats, and recent statistics on insider threats. Further, research objectives, problem statement, justification of the problem, and significance of the proposed work were also deliberated in the chapter. The next chapter reviews the literature on insider threat detection and mitigation.