CHAPTER 2

REVIEW OF LITERATURE

CHAPTER 2

REVIEW OF LITERATURE

2.1. INTRODUCTION

One of the noticeable security challenges in cyber security is insider threat due to their substantial loss of finance and reputation in an organization. It is required to detect and mitigate both intentional and unintentional insider. Thus, a comprehensive literature study on detection and mitigation strategies in the area of insider threat is done in this chapter. This literature study includes the emerging concern of class imbalance problem for its adversity in the precise detection of insider threats. For insider threat detection, the application of machine learning techniques is reviewed. The influence of incorporating biometric analysis to mitigate unintentional insider is studied. This chapter summarizes key findings obtained and research gap identified from a comprehensive review in the area of insider threat.

2.2. SIGNIFICANT REVIEW ON INSIDER THREAT DETECTION

Insider threat is a prominent challenge and the detection of insider threat is required. Existing solution is based on incorporating classification techniques for insider threat detection. The existing research in the area of insider threat detection are discussed in this section.

Asha et al. (2024) examined a unique double-layer architecture to recognize insider threats while handling class imbalance problem. In the first layer, data integration, data transformation for preprocessing, and data sampling is incorporated to handle imbalanced data using eight sampling techniques such as Random Oversampler, SMOTE, ADASYN, Random Undersampler, Edited Nearest Neighbor, Nearmiss1, Nearmiss2, Tomek's link. The well-performing algorithm is identified by evaluating the performance using support vector machine. The balanced data obtained from the first layer is analyzed using seven unsupervised classifiers for insider threat detection, such as one-class support vector machine (OCSVM), iForest, Robust Covariance, KNN, LOF, and LODA. The proposed

architecture with Nearmiss2 and OCSVM achieved f-score and recall of 78.72% and 100%, along with 82.46% accuracy and a higher detection rate for insider threat detection.

Meanwhile, insider threat detection using machine learning techniques such as Decision Tree (DT) and Random Forest (RF) is incorporated (Al-Shehari and Alsowail, 2021). It deals with biased detection results caused by unsuitable encoding techniques and class imbalance problem. It can be overcome by applying feature scaling and a one-hot encoding approach for data encoding. The synthetic minority oversampling technique (SMOTE) is utilized to address imbalanced issues to obtain oversampled data. The data is analyzed to identify users' intense abnormal behavior using six classification algorithms such as Linear Regression, DT, RF, NB, KNN and SVM. It is evaluated using the CERT insider threat dataset. The hybrid solution using a one-hot encoder and SMOTE with DT and RF obtained an AUC of 100%, which is higher than other techniques and successfully handled the class imbalance issue and detection bias insider threat detection. However, above mentioned classification algorithms are prone to mistake negligent unintentional behavior for cautious aberrant actions in both balanced and imbalanced data.

Alternatively, Deep Learning (DL) techniques such as Recurrent Neural Network (RNN), Bidirectional Long Short-Term Memory (BiLSTM), and AutoEncoder (AE) are recommended for superior performance than other machine learning algorithms in conceding malicious activities using imbalanced data.

Alshehri (2022) incorporated a novel approach for insider threat detection through Recurrent Neural Network (RNN) to analyze multivariate time series data containing stream activities. It identifies the relationships among diverse activities to recognize intentional abnormal activities. The performance of RNN is compared with SVM, HMM and shallow neural network (NN), and evaluated using the CERT dataset. While, RNN surpasses other algorithms by obtaining 99% of AUC, 80% of f-score, 99.12% of precision, and 67.12% of recall for recognizing insiders.

Wang et al. (2022) considered the efficient feature selection for achieving an enhanced result, and implemented BiLSTM to select and extract features for analyzing unexpected actions that trigger insider threats in untrusted control devices. In addition,

sliding window for sequential classification is applied to recognize malicious actions among various devices. Initially, six critical features are extracted through generic object-oriented substation events (GOOSE) messages and seven physical features are summarized. The detection accuracy is enhanced using a sliding window algorithm along sequential classification using Bidirectional Long Short-Term Memory (BiLSTM) for insider threat detection and compared with LSTM, SVM, KNN, and DT. BiLSTM obtained a minimal 1% false negative rate and outperforms other algorithms for insider threat detection across diverse devices.

In contrast, Van Ede et al. (2022) designed and employed a novel framework known as DEEPCASE to analyze security activities using deep learning technique such as AutoEncoder. It influences the context containing malicious events and is refined for further scrutiny. It aims to furnish the security events and classify them as malicious, which is abnormal. DEEPCASE approach using the AutoEncoder technique filtered 86.72% of malicious events while reducing physical labor by 90.53% for security officers with a false alarm rate of 0.001%, which is a minimal risk of potential threats.

On the other hand, Folino et al. (2023) detected the abnormal user behavior using elastic stack employing Adaptive Boosting (AdaBoost) based on real-time user activities. It processes and stores daily log information to analyze using ensemble models for classifying user abnormalities in real time by exploiting ELK software architecture in the Kubernetes platform.

Furthermore, a distributed evolutionary technique is adopted to analyze and classify the digital footprints (Folino et al., 2023). The performance of AdaBoost based ELK is compared with OCSVM, LOF, ExtraTrees, gradient boosting, isolation forest, and KNN for real-time insider threat detection. AdaBoost achieves a minimal false alarm rate and an accuracy of 99% while handling missing data in recognizing anomalous user behavior.

Boosting technique is one of the widely applied techniques for insider threat detection. Wei et al. (2022) incorporated Light Gradient Boosting Machine (LighGBM) algorithm for malicious activity detection. Initially, the features are extracted using a bag of words using the IF-IDF model with the aim of constructing a malicious abnormal user

behavior detection framework. Following this, the SEA training group was trained for insider threat detection and obtained an accuracy of 97%, higher than other classification techniques.

On the contrary, for insider threat detection, a hypertuned ensemble algorithm is Gradient Boosting (XGBoost) algorithm is proposed (Mamidanna et al., 2022). It provides an XGBoost framework to address the issue of performance and speed. The performance of XGBoost is enhanced by Grid Search-based hyperparameter optimization that enhances the identification of the micro web using Flask micro web framework in real-time data. The enhanced framework is evaluated using the CERT R4.2 dataset, and the framework attained a maximal accuracy of 98.6% for insider behavior detection.

Instead, Liu et al. (2022) improvised the performance of recognizing real-time atypical actions using optimization techniques such as grid optimization algorithm. It contains multi-modal user and entity behavior analysis (UEBA) framework, namely MUEBA, for analyzing the spatiotemporal representation of data. The framework comprises individual and group analysis for insider threat detection. An LSTM-based individual analysis is accomplished to enhance the sensitivity and augment the performance efficiency of the model. Meanwhile, group analysis is attained through iForest for attribute selection and iTree for construction. The combination of the two models proposed a novel UEBA framework and was evaluated using the CERT R4.2 dataset. The experimental result shows that the combined performance of MUEBA containing LSTM and iForest outperforms the single model of LSTM and iForest based on stability and precision.

Meanwhile, existing research focused on minimizing false alarm rate for insider threat detection using deep learning techniques. AlSlaiman et al. (2023) developed a novel detection framework using a deep learning-based Long Short Term Memory (LSTM) technique to differentiate anomalous activities from genuine ones. The process includes sentiment analysis for classifying user activities and gray encoding of the timely behavior among activities. The framework compares the performance of gray encoding with different data representations containing binary encoding, and real-valued data without encoding for analyzing the timely relationships among behavior. The LSTM-based architecture with gray

encoding enhanced the performance of a 0.29% false positive rate, 97% AUC value, and 2.47% false negative rate using the CERT R4.2 dataset for insider threat detection.

Alternatively, Alajlan and Almasri (2022) optimized a Convolutional Neural Network (CNN) with a multi-population genetic algorithm to obtain an optimized dynamic kernel convolutional neural network for detecting and predicting unusual behavior in a two-phase architecture. The enhanced CNN can differentiate nine diverse attack vectors in the UNSW-NB15 dataset. The performance of optimized CNN shows that it achieved a true positive rate of 99.22%, a false positive rate of 99.11%, a precision of 99.11%, and an F-measure of 99.22% for abnormal insider threat detection with maximum accuracy.

Accordingly, Zhang et al. (2021) proposed an alternate existing solution for insider threat detection using an ensemble learning-based solution along self-supervised algorithm to perform semantic analysis. It incorporates a novel representation of the entity that applies Term Frequency and Inverse Document Frequency (TF-IDF) to examine the significance of the word for further analysis. In addition, the user logs are sampled by applying over-bootstrap with the aim of balancing the target class. Then, the sampled data is analyzed using four deep learning techniques such as CNN, Recurrent Neural Network (RNN), Bagging and Long Short-Term Memory (LSTM) for classifying abnormal activities in insider threat detection. The proposed method is evaluated using CERT R4.2 and R6.2 datasets and achieves an AUC of 99.2% and 95.3% for recognizing malevolent activities.

On the contrary, Al-Mhiqani et al. (2021) proposed a novel framework named AD-DNN that combines the oversampling approach with deep learning technique to handle the class imbalance problem for insider threat detection. The adaptive synthetic (ADASYN) sampling approach is incorporated for oversampling to balance the data distribution. The balanced data is analyzed using the Deep Neural Network (DNN) technique for recognizing insider threats. The performance of the integrated framework is evaluated using the CERT dataset based on accuracy, AUC, F-score, true negative rate, and false positive rate. The performance of ADASYN with DNN increases the total detection rate while obtaining 96% accuracy, 95% f-score, 95% AUC, 96% true negative rate, and 4% false positive rate and outperforms CNN, SVM and LSTM with ADASYN for detecting unusual patterns

On the contrary, a hybrid algorithm was highly employed for describing the existence of unintentional eccentric behavior to suppress misinterpretation levels by applying deep learning algorithms. Accordingly, to handle false alarms and minimal detection rates, the existing study incorporated deep learning algorithms to analyze cyber data streams with changing patterns for detecting intentional insider behavior.

Amuda et al. (2022) highly recommended the techniques comprised of CNN with Gated Recurrent Unit (GRU) to achieve a higher detection rate and evaluated using the CERT dataset. The preprocessing techniques include Min-Max normalization, Standard scaler, and common feature extraction technique incorporated to pick the best performing features for further classification using a proposed hybrid model such as CNN-GRU. The experiment results show that the proposed hybrid model has increased accuracy by 1.48%, precision by 4.21%, and sensitivity by 1.25% for recognizing malicious insider behavior.

Meanwhile, the traditional CNN model is underperforming due to the pooling function and lacks performance for predictive models with spatial correlations. Randive et al. (2023) addressed the above-mentioned challenges by implementing Wavelet CNN with SMOTE. It analyzed features with image representation. Originally, the user log activities are converted into features with single-dimension. It is further converted into images. The enhanced deep learning model incorporates sampling techniques such as SMOTE and ENN for combating class imbalance problem. The combined SMOTE with WCNN obtained the accuracy and AUC of 97.19%, 97.30% for insider threat detection.

Considering a glimpse of diverse anomalous activities in daily life, genetic algorithms is highly effective to optimize the CNN model, and to achieve good performance (Alajlan and Almasri, 2022). Meanwhile, Mamidanna et al. (2022) employed grid search optimization to improve boosting algorithms, namely XGBoost, for real-time insider threat identification.

However, RM and MK (2023) proposed a hybrid solution that incorporates deep learning techniques for preprocessing and normalization from raw data. Subsequently, the features are extracted from normalized data containing seventy-six features and processed using Principal Component Analysis to obtain seven bottlenecks. These features are

processed using soft-clusters of Smart Monkey Optimized Fuzzy C-Means algorithm in a two-way approach. The cluster with attack data is inputted into the Deep learning-based AutoEncoder technique, known for better attack classification. The performance of the proposed approach surpassed other existing approaches by attaining an accuracy of 95% by evaluating the CSE-CIC-IDS-2018 dataset.

Jindal and Singh (2022) incorporated a unique approach for an Intrusion Detection System in a Database based on combined sequential pattern mining with enhanced optimization clustering. It integrates Grey Wolf and Whale optimization algorithms for identifying malicious transactions using Role-based access control and non-RBAC databases. It is considered highly appropriate for spotting unintended anomalous behaviors. A CM-SPADE mining algorithm is applied in database logs to excerpt rules in data dependencies for recognizing outsider threats. Then, user roles are assigned based on proposed metaheuristic clusters for insider threat detection. It identifies malicious transactions that match users' role profiles and compares transaction patterns with gathered dependency rules. The framework is evaluated using the Transaction Processing Performance Council (TPC)-C dataset for detecting malicious transactions and achieved 97.8% accuracy for insider threat detection.

In the area of limiting the issue of false positives, predefined feature engineering techniques are available in existing methods. A multi-fuzzy classifier is more adequate than combined deep learning algorithms for user profiling mechanism (Singh et al., 2022). The preprocessing techniques such as noise reduction, application of isometric feature mapping for feature extraction, and content-based features are utilized for classification. The optimal feature selection based on the emperor penguin technique and multi-fuzzy classifiers for recognizing malicious activities are compared with the Distance method + HMM, deep AE, Fuzzy classifier + Genetic algorithm, LSTM + CNN. It is evaluated using the CERT R4.2 dataset and obtained AUC-ROC of 89.7%, precision of 86.9%, accuracy of 83.3%, recall of 83.9%, detection rate of 89%, f-score of 93%, false positive rate of 2% and error rate of 9%. As a result, false positives are lowered, increased accuracy in threat detection, and enhanced accuracy are obtained.

Al-Shehari and Alsowail (2023) focused on balanced data for malicious activity detection by incorporating DT along with KNN for the slightest false threat detection. It is due to a class imbalance issue in the CERT dataset for insider threat detection, which is overlooked by existing research. A pre-requisite for an effective model aimed at insider threat detection while processing an imbalanced dataset is required. A hybrid data leakage detection model is proposed that applies the variants of random sampling algorithms for handling imbalanced data and various machine learning techniques involving XGBoost, Decision Tree (DT), RF and KNN. The proposed model is evaluated using CERT R4.2 dataset with diverse sampling algorithms. As a result, the performance of Decision Tree and KNN surpasses other algorithms while achieving AUCs of 94% and 87% for enhanced detection of insider threats.

Gayathri et al. (2024) employed a novel algorithm, namely a modified Auxiliary Classifier Generative Adversarial Network referred to as SPCAGAN. It contains a hybrid model using deep learning techniques for analyzing insider threats. The performance of PCA with GAN is compared with other clustering and deep learning algorithms such as multi-layer perceptron, CNN, and Bayesian Neural Network. The proposed SPCAGAN model achieved an average similarity score and silhouette coefficient of 90% and 60% and outperforms other techniques for recognizing insider threats. As a result, the SPCAGAN model obtained a minimal error rate and higher detection accuracy while generating substantial artificial insider activities compared to others.

Al-Shehari et al. (2024a) enhanced an unsupervised algorithm, namely the local outlier factor, to handle the class imbalance problem. It proposed a novel technique, namely the Density-Based Local Outlier Factor (DBLOF) algorithm. The effectiveness of DBLOF for insider threat detection is evaluated using the CERT R4.2 dataset and achieved higher F-score of 98.9%.

Al-Shehari et al. (2024b) incorporated three oversampling techniques, namely SMOTE, Borderline-SMOTE, and ADASYN, to enhance the performance of CNN for insider threat detection while combating the class imbalance problem. The combination of CNN with ADASYN obtained 96% ROC and surpasses the performance of SMOTE and ADASYN. In contrast, a novel Deep Temporal Graph Infomax (DTGI) is proposed to

analyze the user behavior for detecting insider threats using graph neural network (Gao et al., 2025).

Tao et al. (2025) employed enhanced the Test-Train Training model (TTT) by combining Residual Network and Efficient Channel Attention mechanism with Linear layer (TTT-ECA-ResNet). At first, the features are extracted using long-term dependencies. Residual Network and Efficient Channel Attention mechanism is applied to extract the pattern of local features using time-series data. Then, the insider threat is detected using Linear layer. It is evaluated using CERT dataset and achieved 98.75% AUC, and 96.81% F1-score for insider threat detection using TTT-ECA-ResNet.

Eguavoen et al. (2025) employed the hybrid algorithm named HSML-ITD that combines Support Vector Machines (SVM) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for insider threat detection. It applies data preprocessing to handle noisy data, classification using SVM, and predictive learning using ANFIS. It obtained 92% accuracy, 93% precision, 89% recall, and 91% F1-Score using CMU dataset.

Fei et al. (2025) incorporated unsupervised algorithm named LAAEB for insider threat detection that analyses feature information, statistical information, and textual content information in logs for anomaly detection. It applies attention anomaly detection to analyze emails, web page logs. Then, emotion anomaly detection is used to analyze the degree of negative and to find psychological problems. The behavior anomaly detection is used to obtain log embeddings to detect malicious behavior. These results are combined in LAAEB and evaluated using CERT and LANL datasets. LAAEB suppress the FPR by 50% using CERT dataset to obtain 0.06% FPR and 0.97% AUC.

Prasad et al. (2024) enhanced the performance of insider threat detection using a combined framework that integrates sampling techniques, machine learning models, and ensemble learning techniques. Initially, the imbalanced dataset is handled using three oversampling techniques such as Random OverSampling (ROS), SMOTE, and ADASYN and three under sampling techniques such as Random UnderSampling (RUS), Cluster Centroids (CC), and Edited Nearest Neighbors (ENN) to obtain the balanced dataset. The

balanced data is analyzed using five machine learning models namely Logistic Regression, Adaboost, DT, RF, and Naïve Bayes for analyzing the insider pattern. The performance is enhanced by applying ensemble learning techniques. It is known that the combination of ENN, AdaBoost, and PCA obtained the accuracy of 99% for both voting and stacking classifier.

Singh et al. (2023) employed the hybrid algorithm to analyze user behavior for insider threat detection. The features are extracted using Bi-directional Long Short Term Memory (Bi-LSTM), the extracted features are processed using Artificial Neural Network (ANN) for feature selection. The selected features are analyzed using SVM with genetic algorithm to classify the user behavior into genuine and insider. It attained 85.7% precision, 84.4% recall, 84.85% f-score, and 88% accuracy for insider threat detection.

The following table 2.1 shows the review on classification and anomaly detection techniques for intentional insider threat detection.

Table 2.1. Review of Significant Research in Intentional Insider Threat Detection

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
Fei et al. (2025)	Incorporated LAAEB framework that applies attention anomaly detection, emotion anomaly detection , behavior anomaly detection to detect insider threats using text log, and emotion	Classification	LAAEB	It obtains 0.06% FPR with 0.97% AUC for insider detection using CERT and LANL dataset.
Eguavoen et al. (2025)	A framework named HSML-ITD is proposed for insider threat detection using SVM and ANFIS	Classification	SVM and ANFIS	It attained enhanced performance by obtaining 92% accuracy, 93% precision, 89% recall, and 91% F1-Score.

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
Tao et al. (2025)	Proposed TTT-ECA-ResNet model that combines Test-Train Training model for feature extraction, Efficient Channel Attention to extract local feature pattern, ResNet with Linear layer for insider threat detection using CMU dataset	Classification	TTT-ECA-ResNet	It obtained 98.75% AUC, and 96.81% F1-score for and outperforms other existing techniques.
Gao et al. (2025)	A Deep Temporal Graph Infomax with graph network and anomalous sample generator is proposed for insider detection.	Classification	DTGI, RShield	DTGI obtained 95.52% accuracy, 96.79% recall, 95.58% F1-score, and 96.32% AUC.
Al-Shehari et al. (2024a)	Employed a novel technique namely DBLOF for insider threat detection based on detection rate and f-score.	Anomaly detection	DBLOF	It obtained a detection rate and f-score of 98% and 98.9% using the CERT dataset for handling class imbalance problem.
Al-Shehari et al. (2024b)	Class imbalance problem in CNN is handled using SMOTE, ADASYN and Borderline-SMOTE.	Classification	CNN	CNN with ADASYN achieved higher performance than SMOTE and Borderline SMOTE with 96% ROC.
Prasad et al. (2024)	Integrated the ENN for handling imbalanced data, AdaBoost for classification, PCA for dimensionality	Classification	ROS, SMOTE, ADASYN, RUS, CC,	It obtained the accuracy of 99% and outperforms other sampling methods and machine learning

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
	reduction and applied ensemble learning to enhance the performance of insider threat detection.		ENN, Logistic Regression, Adaboost, DT, RF, and Naïve Bayes	models for insider threat detection.
Singh et al. (2023)	Integrated Bi-LSTM and ANN for feature extraction and feature selection. Applied SVM and GA for classification for insider threat detection	Classification	SVM and GA	SVM and GA obtained the accuracy of 88% with 85.7% precision, 84.4% recall, and 84.85% f-score using CERT R4.2 dataset.
Folino et al. (2023)	Explored elastic stack (ELK) to distinguish user behavior and ascertain anomalies in real-time daily logs using ensemble models in the Kubernetes platform	Anomaly detection – Boosting	Adaboost, OCSVM, LOF, ExtraTrees, Gradient boosting, IF, and K-NN	Achieved lower false alarm rate with 99% accuracy.
Asha et al. (2023)	An architecture containing two layers is examined for MIT detection involving sampling technique, namely along the OCSVM classifier to depreciate data disparity.	Classification – Deep Learning	OCSVM, IF, Robust Covariance, KNN, LOF, LODA	OCSVM surpasses recall, f-score, and accuracy with 100%, 78.72%, and 82.46%.
AlSlaiman et al. (2023)	Developed a novel architecture to differentiate abnormal activities in daily life by applying LSTM.	Classification – Deep Learning	LSTM	Enhanced performance is achieved in terms of false positive rate, false negative rate, and AUC is 0.29%, 2.47%, and 97%.

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
Liu et al. (2023)	Proposed a multi-modal framework containing individual and group analysis using LSTM and iForest algorithm.	Classification – Hybrid algorithm	LSTM, iForest	Combining LSTM and iForest outperforms single LSTM and iForest in terms of stability and precision.
Randive et al. (2023)	Developed and combined SMOTE sampling technique with Wavelet Convolutional Neural Network (WCNN) to deal with class imbalance problem using image representation for insider threat detection.	Classification -Hybrid	WCNN	Obtained an accuracy of 97.19% and AUC of 97.30% with a low positive rate for malicious insider threat detection.
RM and MK (2023)	Proposed a hybrid deep learning approach using smart monkey optimized fuzzy C-Means algorithm (SMO-FCM) and AE for insider threat detection.	Classification -Hybrid	SMO-FCM, AE	Outperform other state-of-the-art methods in terms of accuracy by over 95%.
Amuda et al. (2022)	A hybrid algorithm that demoted high false alarms and fewer detection rates were proposed to accommodate uncertain patterns in structured network data streams to detect malicious insider activities using imbalanced data.	Classification – Hybrid algorithm	CNN and GRU	The proposed model enhances the accuracy, precision and sensitivity rate to 1.48%, 4.21%, and 1.25% than existing methods using a CERT dataset.

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
Gayathri et al. (2022)	Applied a unique hybrid model for analyzing insider threat with obtained balanced data using a modified SPCAGAN using balanced data.	Clustering	MLP, CNN and Bayesian Neural Network	Evaluated using PCA and Obtained an average of 90% and 60% in terms of Similarity Score and Silhouette Coefficient using a hybrid algorithm and outperformed others for insider threat recognizing.
Alshehri (2022)	Developed a novel model for analyzing multiple audit reports using RNN using imbalanced data to learn the relationships between various activity streams for recognizing malicious activity.	Classification – Deep learning	SVM, HMM, shallow Neural Network (NN), RNN	RNN outperforms others in terms of AUC, f-score, precision, and recall is 99%, 80%, 99.12%, and 67.12%.
Wang et al. (2022)	Enhanced the performance of BiLSTM networks by selecting six essential features and applying a window algorithm for splitting them into window snippets using imbalanced data.	Anomaly detection – Deep Learning	BiLSTM, LSTM, SVM, KNN, and DT	Obtained a minimum false-negative rate of 1% and outperformed other ML algorithms.
Van Ede et al. (2022)	Designed and evaluated a novel architecture named DEEPCASE using imbalanced data that	Classification – Deep Learning	AE	Filtered 86.72% of malicious events, while the false alarm rate is minimal at 0.001%.

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
	analyses every activity to recognize an abnormal one.			
Jindal & Singh (2022)	Combining Grey Wolf and Whale optimization algorithms into hybrid clustering to identify malicious from normal behavior in imbalanced data using the TPC-C dataset.	Clustering – Optimization	Grey Wolf and Whale optimization algorithm	Obtained an accuracy of 97.8% for recognizing insider threats.
Singh et al. (2022)	Explored multi-fuzzy classifier to parallelly recognize malicious activity based on daily activities using CMU-CERT version 4.2.	Classification – Hybrid algorithm	Distance method + HMM, deep AE, Fuzzy classifier + Genetic algorithm, LSTM + CNN, multi fuzzy classifier	AUC-ROC, precision, accuracy, recall, detection rate, f- score, false positive rate, and an error rate of 89.7%, 86.9%, 83.3%, 83.9%, 89%, 93%, 2% and 9%.
Alajlan & Almasri (2022)	Optimized a CNN with a multi-population genetic algorithm for differentiating nine different types of attacks in a two-stage architecture, including malicious detection and prediction scheme.	Classification – Deep Learning	CNN	Evaluated using the UNSW-NB15 dataset based on true positive rate, false positive rate, precision, and f-score are 99.22%, 99.11%, 99.11%, and 99.22%.

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
Mamidanna et al. (2022)	Hypertuned an ensemble algorithm, namely XGBoost, using Grid Search optimization enhanced micro web identification model using real-time data.	Classification – Boosting	XGBoost	Generated a higher accuracy of 98.6% than other state-of-art methods.
Wei et al., (2022)	Constructed malicious activity detection framework applying LightGBM algorithm.	Classification – Boosting	LightGBM	Obtained satisfying accuracy of 97%.
Al-Shehari & Alsowail (2022)	Recommended a hybrid insider detection model applying variants of random sampling algorithms and machine learning methods to handle extreme data discrimination in CERT data.	Classification – Hybrid solution	XGB, DT, RF and KNN	DT and KNN outperform others, obtaining 94% and 87% of AUCs.
Zhang et al. (2021)	Explored ensemble learning along a novel entity representation method named TF-IDF and applied over-bootstrap sampled user logs to balance the data for analyzing insider threat detection.	Classification – Deep Learning	CNN, RNN, Bagging, LSTM	Obtained AUCs for CERT4.2 and CERT6.2 datasets are 99.2% and 95.3%.
Al-Mhiqani	Solved imbalanced problem in DNN by applying	Classification – Deep	CNN, SVM, DNN and	AD-DNN obtained outperforms others in

Study	Key findings	Machine learning model scheme	Algorithms applied	Observations
et al. (2021)	ADASYN to balance unequal data distribution for insider threat detection.	Learning	LSTM	terms of accuracy, f-score, AUC, true negative rate and false positive rate with 96%, 95%, 95%, 96%, and 4%.
Al-Shehari & Alsowail (2021)	Explored various machine learning techniques and applied SMOTE, one-hot and label encoding to depress misclassification while analyzing user behavior in balanced data.	Classification – Hybrid solution	Linear Regression, DT, RF, NB, KNN and SVM	DT and RF suppress others in terms of AUC with 100%.

From the above table 2.1, it is observed that existing studies focus on detecting insiders which is intentional but fails to further classify the insider into intentional or unintentional. Next to review on intentional insider detection is review on unintentional insider detection.

Unintentional insider detection

Existing research addresses the challenge of unintentional insider threats using behavior analysis and policies. Some of research on unintentional insider detection is reviewed in this section.

Rahman et al. (2022) developed a framework named NARX that studies the threat landscape of unintentional insiders in terms of threat level (such as access, motivation and action) and insider level (such as +action, method and knowledge). The Autoregressive Integrated Moving Average (ARIMA) is used to train the NARX framework to perform time series analysis. Neural network is used for evaluating the performance of unintentional

insider prediction. It obtained 97% detection rate for unintentional insider detection using NARX framework.

Khan et al. (2022) investigated the Critical Decision Method (CDM) for analyzing the unintentional insiders using grounded theory. Data is gathered by interviewing ten individuals who experienced unintentional cybersecurity breaches. A five-pillar action plan was developed based on decision making, task factors, accidents, and organisational factors. It is analyzed that organization factor plays major role followed by decision making, task factors and accidents.

The following table 2.2 shows the review of significant research in unintentional insider detection.

Table 2.2. Review of Significant Research in Unintentional Insider Threat Detection

Study	Key findings	Scheme	Algorithms applied	Observations
Rahman et al. (2022)	Developed NARX framework based on threat level and insider level for unintentional insider detection	Behavior Analysis	ARIMA, Neural network	Obtained detection rate (97%) using CERT dataset
Khan et al. (2022)	Critical Decision Method (CDM) is developed to analyze the unintentional insiders using grounded theory approach.	Policy	-	A five pillar action plan as output is developed.

From the table 2.2, it is observed that very few evident study available and focuses on unintentional insider detection compared to intentional insider detection.

From the table 2.1 and 2.2, it is observed that both intentional and unintentional insider detection is prone to class imbalance problem and it need to be handled.

Class Imbalance Problem

Class imbalance problem is a major issue while detecting intentional and unintentional insider threats. Since the instances of insider activity are less compared to instances of genuine user behavior, which results in misclassification and misinterpretation

due to the negligence of minority instances. At present, the solution for insider threat detection using imbalanced data is inappropriate. Thus, selecting the best detection approach is still required. The below section discusses the existing solutions for the class imbalance problem available so far.

Janjua et al. (2020) explored six ML techniques including AdaBoost, Naïve Bayes, Linear Regression, Logistic Regression, K-Nearest Neighbor, and Support Vector Machine for recognizing the malicious insiders for choosing the algorithm that aims at achieving best performance and evaluated based on accuracy, Area Under Curve (AUC) and recall using The Wolf of Sutd (TWOS) dataset. Adaboost performed well and obtained the accuracy, recall and AUC of 98.3%, 98% and 98.3%, outperforming other techniques for malicious emails using the TWOS dataset.

Jiang et al. (2018) implemented user behavior analysis for insider detection using Extreme Gradient Boosting (XGBoost), Random Forest, and Multi-layer Perceptron and evaluated using accuracy, precision, recall, and f-score. As a result, XGBoost performed better than other algorithms and obtained a 99.96% f-score, 95.54% accuracy, 95.54% precision, and 100% recall using the Computer Emergency Response Team (CERT) dataset for user behavior analysis for the role of Electrical Engineering compared with the salesman and IT admin.

Boosting techniques involving a Light Gradient Boosting Machine (lightGBM) have been incorporated into the intelligent framework for insider detection (Mohammed et al., 2021). It was evaluated using f-score, AUC, and accuracy. The classifier achieved an accuracy of 99.47%, AUC of 99.79%, and F1-score of 92.26% using imbalanced CERT data and successfully obtained higher AUC and f-score for detecting the insider threat event. It successfully handled the Class Imbalance Problem compared to other state-of-art methods.

Pantelidis et al. (2021) concentrated on the highly imbalanced malicious insider data. Deep learning techniques such as AutoEncoder (AE) and Variational AutoEncoder (VAE) were incorporated and evaluated using performance metrics such as precision, recall, accuracy, and f1-score for insider detection. The performance of the VAE neural network

provides the higher performance by obtaining precision, recall, accuracy and f1-score of 92%, 96%, 96% and 94%, and outperforms AE.

Le and Zincir-Heywood (2018) evaluated three ML algorithms, namely Decision Tree (DT), Hidden Markov Model (HMM), and Self Organizing Maps (SOM), to choose the outperforming algorithm with the aim of classifying malicious activities based on performance criteria such as recall, FPR, and accuracy. The performance of VAE and SOM achieved a maximum detection rate with less False Positive Rate (FPR). However, SOM provides better results based on detection rate, false-positive rate, and accuracy in detecting the insider threat using CERT data.

Insider threat detection using an imbalanced dataset is quite cumbersome, and the CIP makes it hard to achieve better performance. It can be solved by applying a hybrid solution to the issue of class disparity in insider anomaly detection. Sheykhkanloo and Hall (2020) introduced a framework using the spread subsample feature to curb CIP in five supervised machine learning techniques such as Naïve Bayes, Linear Regression, Random Forest, Support vector machine, and Neural Network. It was evaluated based on the performance of five classifiers for insider detection using precision, recall, f-score, and time taken. The diverse parameters in the classifier influence their performance and are measured using performance metrics. As a result, compared to the balanced dataset, its impact is more significant on the imbalanced dataset. It is observed that SVM and Naïve Bayes (NB) perform well compared to other machine learning techniques.

The performance can be further achieved by applying boosting algorithms. Chattopadhyay et al. (2018) suppressed the class imbalance problem using the Synthetic Minority Over-sampling Technique (SMOTE), an oversampling technique to equalize the data proportion. The performance is evaluated using three-time series classification methods, namely Random Forest (RF), Isolation Forest (IF), and deep Auto Encoder (AE) based on performance metrics such as precision, recall, and f-score. Thus, deep AE and RF worked better using balanced data.

Al-Shehari and Alsowail (2021) proposed a framework based on machine learning techniques for insider threat detection that handles class imbalance problem. In preprocessing, it involves feature scaling, one hot encoding technique and synthetic minority oversampling technique (SMOTE) using the CERT dataset and evaluated using

precision, recall, f-score and AUC. SMOTE is useful in equalizing the class data and evaluated using six machine learning classifiers, namely Logistic Regression, Random Forest, Naïve Bayes, K-Nearest Neighbor (KNN), Decision Tree, and Kernel-SVM (KSVM). In contrast, Decision Tree achieved AUC-ROC up to 100% compared to Logistic Regression, Random Forest, Naïve Bayes, and KNN with AUC-ROC of 79%, 100%, 84%, and 99% and outperformed these techniques.

Apart from SMOTE, Adaptive Synthetic (ADASYN) is highly recommended with Deep Neural Network (DNN) to deal with class imbalance problem in insider threat detection (Al-Shehari & Alsowail, 2021). The performance of AD-DNN is compared with SVM, LSTM, DNN, OCSVM-based DBN, and LSTM-Autoencoder based on Accuracy, F-score, AUC, FPR, FNR and TNR. However, AD-DNN obtained 96% accuracy, 95% f-score, 95% AUC, 4% FPR, 5% FNR, and 96% TNR. It outperforms other techniques for insider threat detection.

Noever (2018) explored 88 machine learning algorithms to select the outperforming technique based on evaluation metrics such as kappa, time and accuracy. Some of the higher performance attained machine learning algorithms include Bayesian model, linear regression, logistic regression, model tree, linear classifier, neural network, RF, SVM, Tree based model, Rule based model, Partial least squares and polynomial model for insider threat detection. In contrast, Random Forest achieved accuracy equal to 98% and outperforms other techniques based on above mentioned evaluation metrics. Furthermore, it is complicated to detect the new malicious intentional insider threat using classification algorithms; therefore, anomaly detection techniques are explored.

On the other hand, anomaly detection is widely incorporated for insider threat detection using unsupervised ensemble learning algorithms. Le and Zincir-Heywood (2021) applied AutoEncoder, Isolation Forest, Local Outlier Factor (LOF), and Lightweight On-line Detector of Anomalies (LODA) for analyzing temporal information. The experimental result shows that Autoencoder outperforms other algorithms and state-of-the-art approaches in voting metrics by obtaining a high detection rate while preserving the least false positive rates.

The characteristics of spatial and temporal information of user behavior are analyzed for recognizing insider threats. Ferreira et al. (2019) incorporated zero-knowledge anomaly-based framework comprising feature normalization techniques such as a standard scaler, maximum absolute value scaler, quantile transform. Machine learning-based temporal information analysis is accomplished using Random Forest, Logistic Regression, and Multi-layer perceptron. The experimental result shows that the performance of the Standard scaler with Random Forest obtained higher f-score, precision, and recall compared to other algorithms for insider threat detection.

Jiang et al. (2019) introduced the application of the anomaly detection model using a Graph Convolutional Network (GCN) to recognize both individuals' abnormal behavior and accompanying abnormal groups in the structural information of graphs. It evaluated the performance with other algorithms, namely Random Forest, Logistic Regression, Support Vector Machine and Convolutional Neural Network (CNN), using accuracy, precision and recall. In comparison, GCN obtained a recall of 83.3%, accuracy of 94.5%, and precision of 100%, which outperforms other algorithms for fraud activity detection.

A framework based on ensembles of deep auto encoder was highly recommended for insider threat detection and trained each Autoencoder with specific audit data containing precise, genuine behavior (Liu et al., 2018). The abnormal and genuine behavior is distinguished by calculating reconstruction error for original and decoded data. Subsequently, a joint decision-making mechanism is proposed to identify the total genuine (non-malicious) score. The performance shows that deep AE recognize every malicious activity along minimal false positive rate.

Insider threat is a major concern of research in recent years. However, merely countable hybrid solutions are presented. Al-Mhiqani et al. (2022) introduced a multi-layer framework involving two layers where the high performing classification model is chosen for insider threat detection using entropy-VIKOR based multi-criteria decision making technique in layer-one. In the subsequent layer, the hybrid model for insider threat detection combining Misuse Insider Threat Detection using Random Forest and anomaly insider threat detection using K-Nearest Neighbors was developed and recommended for detecting familiar and unfamiliar insider threats. The framework is evaluated using the CERT R4.2

dataset in terms of Recall, Accuracy, F-score, Precision, False Positive Rate (FPR), False Negative Rate (FNR), AUC, and True Negative Rate (TNR). The performance of the framework is evaluated using the CERT R4.2 dataset and attained the maximum performance for detecting intentional insider threats of 99% overall accuracy, 0.29% FPR with a train-test ratio of 80%-20% for recognizing unintentional insider threats of 97% accuracy, 2.88% FPR.

Alternatively, insider threats are recognized by analyzing user emailing activities in the proposed hybrid model (Garba et al., 2021). It combines Principal Component Analysis (PCA) and k-means for anomaly detection and is evaluated using an imbalanced CERT R6.2 dataset. Preprocessing is accomplished using natural language processing techniques that involve topic modelling to procreate vector space for training anomaly detection techniques with the aim of detecting malicious content in an email. The proposed framework combining PCA and K-means has obtained an 89% detection rate and outperforms the K-means and PCA in terms of anomalous values of 1%, 5%, 10%, 15%, 20%, 25%, and 30% cut-off. However, the performance was not satisfactory.

Diop et al. (2019) combined an intelligent framework using classification techniques and graph methods such as linear algebra and parallel computing algorithms for insider threat detection by analyzing user and entity behavior using an anomaly detection algorithm. Nine classifiers involving isolation forest, one-class support vector machine, local outlier factor, elliptic envelope, artificial neural network, Gaussian naive Bayes (Gnb), ensemble learning Bagging classifiers (Bgc), random forest and gradient boosting (Gbc) were applied and evaluated based on f-score, AUC, precision and recall. However, boosting techniques such as Bgc, RF, and Gbc obtained a maximum detection rate with F-score, recall, precision, and AUC of 99%, surpassing other classifiers.

In contrast, Le and Zincir-Heywood (2020) applied an unsupervised machine learning algorithm for a hybrid ensemble-based anomaly detection approach to recognize insider threats. It involves Autoencoder and isolation forest with diverse temporal data representation approaches such as concatenation, percentile, and mean difference. It analyzed the temporal representation of data using AE and IF and evaluated to detect employee behaviour change in accordance with adversarial constraints to achieve maximal

detection rate and slightest false positive rate. Thus, integrating percentile representation with AE as a hybrid approach attained the higher performance and surpassed Isolation Forest for detecting insider threats.

In addition, the hybrid approach in deep neural networks for anomaly detection is beneficial for detecting insider threats (Yuan et al., 2020). The framework combines Long-Short Term Memory (LSTM) and CNN for behavior modeling. LSTM trains the user activities to extract the features containing condensed temporal information; CNN subsequently converts the condensed features into a definite feature matrix for insider threat detection. As a result, the performance of combined LSTM-CNN obtained AUC up to 94.49% using an imbalanced CERT dataset.

In the case of handling class imbalance problem and detection bias, Al-Shehari and Alsowail (2021) applied a synthetic minority oversampling technique (SMOTE) following various encoding approaches involving feature scaling and a one-hot encoder. However, a novel sampling technique, Conditional Generative Adversarial Network (CGAN), was proposed and compared with two oversampling techniques, ROS and SMOTE (Gayathri et al., 2021). The performance of CGAN is evaluated using four ML techniques, namely RF, XGBoost, MLP, and Intelligent Deep Convolution Neural Network (IDCNN) based on precision, recall, f-score, kappa, and Mathews Correlation Coefficient (MCC). The performance of CGAN compromised ROS and SMOTE by achieving 0.99 AUC-ROC value for multi-class classification for insider threat detection.

However, Al-Shehari et al. (2023) evaluated the performance of anomaly-based Isolation Forest algorithm with Extreme Gradient Boosting, Random Forest, Decision Tree and KNN based on accuracy, f-score and detection rate. Comparatively, random forest attained higher performance with an accuracy and detection rate of 98% and an f-score of 99% for insider threat detection using the CERT dataset.

Besnaci et al. (2023) proposed a novel machine learning model, namely S-LSTM that combines the SMOTE oversampling technique with LSTM to handle the class imbalance problem. The performance of S-LSTM is enhanced by obtaining 99% accuracy using the CERT R4.2 dataset for insider threat detection. On the other hand, Pal et al.

(2023) effectively handled the class imbalance problem by employing the Equally Weighted Random Sampling (EWRS) technique. The balanced data is analyzed using a stacked ensemble of attention models, such as stacked-LSTM and stacked-GRU-based attention models for insider threat detection. It attained a 0.99 AUC score on CERT R4.2 and R5.2 and obtained an AUC of 0.97 on R6.2.

Jaiswal et al. (2024) proposed a Two-Step Insider Threat Detection (TSITD) approach where preprocessing is accomplished in the first layer. In second layer, TSITD, a combination of sampling techniques such as SMOTEENN (SMOTE and ENN), SMOTETomek (SMOTE-TL), ADASYN, SVM+SMOTE, ENN and classifiers such as ANN, Random Vector Functional Link (RVFL), Extreme Learning Machine (ELM), RF, DT, XGBoost, kNN, and SVM is employed to handle class imbalance problem.

Gao et al. (2025) detected the insider threat using a novel method namely Deep Temporal Graph Infomax (DTGI) that contains heterogeneous graph network and anomalous sample generator to generate attack samples. DTGI obtained the AUC of 0.9632 and outperforms RShield and other existing methods for insider threat detection using CERT dataset.

The following Table 2.3 reviews existing state-of-the-art methods on the class imbalance problem.

Table 2.3. Review of Significant Research in Class Imbalance Problem

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
Gao et al. (2025)	Detection using imbalanced data	Introduced DTGI for insider detection with heterogeneous graph network and anomalous sample generator.	Classification	DTGI, RShield	DTGI outperforms RShield by obtaining 95.52% accuracy, 96.79% recall, 95.58% F1-score, and 96.32% AUC.

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
Jaiswal et al. (2024)	Detection using imbalanced data	Proposed a Two-Step Insider Threat Detection approach while handling class imbalance problem using SMOTEENN, SMOTETomek, ADASYN, SVM+SMOTE, and ENN. Evaluated its performance using various classifiers.	Classification	ANN, RVFL, ELM, RF, DT, XGBoost, kNN, SVM	TSTID model achieves better performance in terms of accuracy, precision, recall and f-score.
Al-Shehari et al. (2023)	Detection using imbalanced data	Evaluated five classifiers to select the best performing technique for insider threat detection based on accuracy, detection rate and f-score.	Anomaly detection	IF, Extreme Gradient Boosting, RF, DT and KNN	Anomaly detection using Isolation Forest performed well compared to others by obtaining accuracy, detection rate, and f-score of 98%, 98%, and 99%.
Besnaci et al. (2023)	Detection using balanced data	Proposed S-LSTM model integrating the SMOTE technique for generating synthetic samples to balance the dataset.	Classification	LSTM	Obtained higher prediction accuracy of 99% and effective in detecting insider threats.

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
Pal et al. (2023)	Detection using balanced data	Introduced EWRS technique for handling data imbalance. Detected the insider threat using a stacked ensemble of attention models.	Classification	stacked-LSTM and stacked-GRU-based attention models	Achieved an average AUC score of 0.99 on CERT v4.2 and v5.2, and 0.97 on v6.2.
Al-Mhiqani et al. (2022)	Detection in the imbalanced dataset	Proposed a multi-layer framework using misuse insider threat detection and anomaly insider threat detection to detect known and unknown insider threats.	Anomaly detection	KNN and RF	The performance of the proposed hybrid model obtained accuracy and FPR of 99% and 29% and outperformed other state-of-the-art methods.
Garba et al. (2021)	Detection in imbalanced data	Proposed a framework based on behavior of malicious insiders that combines the k-means and PCA for anomaly detection using an imbalanced CERT r6.2 dataset.	Anomaly detection	K-means, PCA	The proposed framework obtained a detection rate of 89% and outperforms the K-means and PCA based on cut-off values of 1%, 5%, 10%, 15%, 20%, 25%, and 30% cut-off.
Gayathri et al. (2021)	Detection in the balanced dataset	Proposed CGAN and evaluated three oversampling methods with four anomaly detection techniques for CIP.	Anomaly detection	RF, XGBoost, MLP and IDCNN	The anomaly detection using CGAN in four classifiers has obtained higher performance than ROS and SMOTE to

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
					perform multi-class classification.
Mohammed et al. (2021)	Detection in imbalanced data	Introduced an intelligent framework for insider threat detection using LightGBM and evaluated using f-score, AUC and accuracy.	Classification	LightGBM	Achieved an accuracy of 99.47% using imbalanced cert data and successfully obtained higher AUC and f-score for detecting the insider threat event.
Pantelidis et al. (2021)	Detection in imbalanced data	Concentrated on the highly imbalanced malicious insider data, the AE and VAE were evaluated using performance metrics such as precision, recall, accuracy, and f1-score.	Classification	AE and VAE	The performance of the VAE neural network provides the best result and obtained precision, recall, accuracy and f1-score of 92%, 96%, 96% and 94%, which is higher than AE and outperforms AE.
Al-Shehari, Alsowail (2021)	Detection in the balanced dataset	Applied ADASYN for CIP and evaluated using DNN to detect the insider threat.	Classification	Deep neural network (DNN)	The performance of DNN outperforms existing ML techniques using balanced data.
Le & Zincir-Heywood (2020)	Detection in imbalanced data	Evaluated two unsupervised ensemble-based anomaly detection techniques using concatenation, percentile and mean	Anomaly detection	AE and IF	The combination of percentile representation of data in AE outperforms IF and achieves a high detection rate with a low

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
		difference, to describe the user behavior changes.			false-positive rate to detect insider threats.
Janjua et al. (2020)	Detection using imbalanced data	Evaluated six ML classifiers using the TWOS dataset to select the best performing technique for insider threat detection based on accuracy, recall and AUC.	Classification and regression	Adaboost, NB, LR, KNN, LR and SVM	Adaboost performed well and obtained the accuracy, recall and AUC of 98.3%, 98% and 98.3%, outperforming other techniques for malicious emails using the TWOS dataset.
Sheykhkanloo & Hall (2020)	Detection using imbalanced data	A framework was proposed using supervised methods to detect the insider threats and evaluated using balanced data via the Spread Subsample feature in the Weka tool and imbalanced data in five classifiers based on precision, recall, f-score, and time taken.	Classification	NB,LR, RF, SVM, and NNs	The classifiers with different parameters affected the performance metrics, but the impact was more substantial on the imbalanced dataset than on the balanced dataset.
Noever (2019)	Detection in imbalanced data	Applied different ML techniques for successful insider threat detection and evaluated using	Classification	Bayesian model, LR, logistic regression, model tree,	Random forest obtained an accuracy of 98% and performed better than other classifiers.

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
		accuracy, kappa and time.		linear classifier, neural network, RF, SVM, Tree based model, Rule based model, Partial least squares, polynomial model.	
Le DC, Zincir-Heywood (2019)	Detection using imbalanced data	Proposed novel insider threat detection framework using anomaly detection in an unsupervised ensemble approach using the CERT dataset	Anomaly detection	AE, IF, LODA and LOF	AE outperforms the other algorithms based on voting metrics to detect the insider threats.
Ferreira et al. (2019)	Detection in imbalanced data	A framework was proposed to model the zero-knowledge anomaly-based behavior for insider threat detection.	Anomaly detection	LR, RF, and MLP	The performance of RF produces the best result for detecting insider threats than other techniques.
Jiang et al. (2019)	Detection in imbalanced data	Designed the anomaly detection using a graphical neural network, namely	Anomaly detection	RF, SVM, LR, CNN and GCN	On the basis of accuracy, precision, and recall, GCN surpasses other algorithms.

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
		GCN, to detect the abnormal malicious behavior and evaluate using other existing methods.			
Diop et al. (2019)	Detection in imbalanced data	Combined an intelligent framework using classification and graph methods for insider threat detection. Evaluated using f-score, AUC, precision and recall.	Anomaly detection	IF, OCSVM, LOF, EE, ANN, Gnb, Bgc, RF and Gbc	The boosting techniques obtained f-score, recall, precision and AUC scores of 99% and outperformed other algorithms.
Le DC & Zincir-Heywood (2018)	Detection in the imbalanced dataset	Evaluated three ML algorithms to select the best algorithm to classify malicious behavior based on recall, FPR and accuracy.	Classification	SOM, HMM, and DT	SOM provides better results based on detection rate, false-positive rate, and accuracy in detecting insider threats using CERT data.
Jiang et al. (2018)	Detection in imbalanced data	Implemented user behavior analysis for MIT detection using XGBoost and evaluated using accuracy, precision, recall, and f-score.	Classification	RF, MLP and XGBoost	XGBoost performed better than other algorithms and obtained a 99.96% f-score using the CERT dataset for user behavior analysis.
Chattopadhyay et al. (2018)	Detection in balanced data	Applied oversampling technique, namely SMOTE to handle	Classification	IF, RF, and Deep AE	The performance of RF and Deep AE is comparable and

Study	Study scheme	Key findings	ML model scheme	Algorithms applied	Observations
		CIP and evaluated classification methods.			achieves the best result for detecting insider threats.
Liu et al. (2018)	Detection in the imbalanced dataset	Proposed an ensemble of deep AE for anomaly detection, trained using normal and abnormal behavior, and evaluated based on accuracy.	Anomaly detection	Deep AE	Deep AE detects any malicious insider behavior with a minor FPR.
Yuan et al. (2018)	Detection in the imbalanced dataset	Developed a double-layer framework using LSTM and CNN and evaluated using AUC.	Anomaly detection	LSTM and CNN	LSTM and CNN obtained AUC = 0.9449, and detects maximum insider threat in CERT data.

From the above table 2.3, it is observed that most existing studies have solved the class imbalance problem in insider threat detection using sampling techniques. Very few tuned the sampling techniques to handle class imbalance problem. Next to the class imbalance problem is review on insider threat mitigation, which is discussed in the next section.

2.3. SIGNIFICANT REVIEW ON INSIDER THREAT MITIGATION

Only employing insider threat detection is not effective due to extensive financial and reputation loss. Thus, the substantial mitigation of the insider threat is required. User authentication is one of the strategies for insider threat mitigation (Sumitra et al., 2014). User authentication using behavioral biometrics is used to examine the uniqueness of individuals to accomplish certain activities. The following figure 2.1 shows the different categories of biometrics available so far.

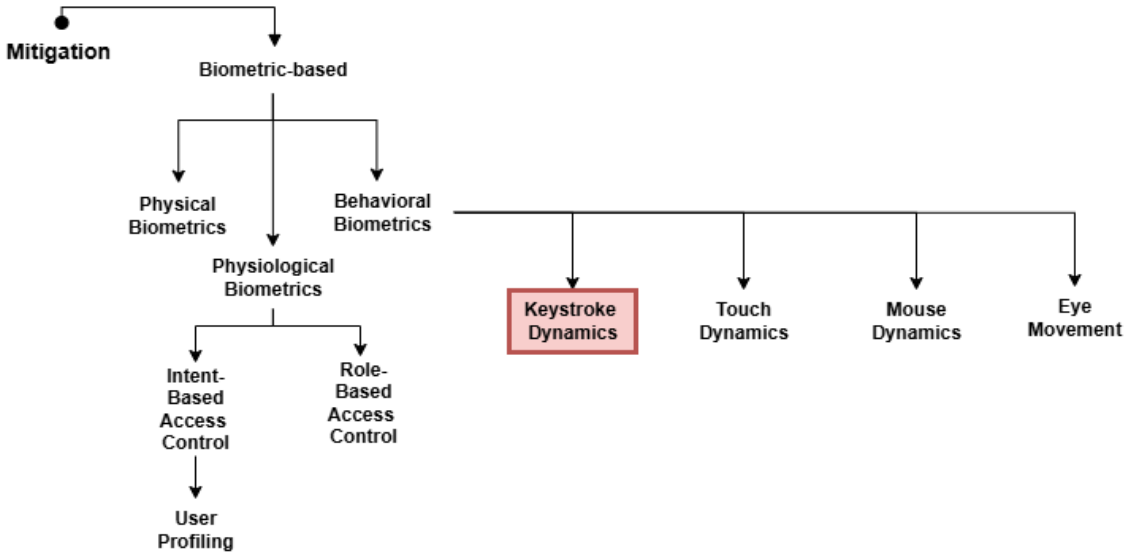

Figure 2.1 Categories of Biometric Techniques

This research focuses on keystroke dynamics for user authentication for insider mitigation. This section discusses various existing methods available so far in user authentication.

The existing solutions highly recommend the utilization of deep learning techniques for user authentication using keystroke dynamics. However, the performance of deep learning techniques requires optimization, which is accomplished by fine-tuning the required parameters for performing classification.

Tewari and Verma (2022) incorporated Support Vector Machine for extracting temporal features of keystrokes for accomplishing pre-trained model. The model integrates PCA with SVM for enhanced feature selection for extracting the selected features and is converted into image format using the CNN technique containing two pre-trained models, Resnet and Alexnet, for biometric authentication. The experimental result shows that the performance of ResNet surpasses AlexNet by obtaining an accuracy of 98.57% with selected features from PCA+SVM.

One of the major characteristics of keystroke dynamics is to analyze the information regarding behavior and relevant context, focusing on user authorization by distinguishing abnormal typing patterns from their unique, genuine behavior.

In case of integrating electromagnetic-triboelectric nanogenerators for handling password vulnerabilities, Maharjan et al. (2021) implemented hybrid sensors to gather common passwords in the form of key mechanical energy where electrical signals are retrieved. Meanwhile, the Artificial Neural Network technique processes these signals to obtain the trained model for establishing a biometric authentication system to identify and authenticate the user. As a result, the performance of ANN is compared with SVM, where ANN attained an accuracy of 99% and surpasses SVM for biometric authentication by incorporating hybrid sensors.

Meanwhile, Wang et al. (2022) proposed a user authentication framework named SIURUA to analyze the features calculated based on irrelated scene and user-related behavior for user identification by implementing a combination of keystroke and mouse dynamics. At initial, the features related to keystrokes and mouse behavior are extracted. Subsequently, features irrelevant to scenes containing minimal correlation are acquired. In addition, the features obtained from irrelevant scenes are integrated with user-relevant information to achieve data integration. The fused data is classified using a Support Vector Machine for analyzing user behavior. The SIURUA framework containing SVM attained an accuracy of 84% and outperformed other user authentication methods by analyzing information regarding keystrokes and mouse movement.

In contrast, past research suggests enrichment in features engineering for better performance of keystroke authentication using Machine learning algorithms. In particular, Kim et al. (2020) introduced an innovative feature selection technique based on a filter algorithm using a feature scoring approach to authenticate keystroke dynamics. Parameters such as trimmed mean and coefficient of variation are utilized to compute novel feature selection techniques. It incorporated multi-factor PIN-based authentication for keystroke authentication and was compared with past feature selection algorithms. It is evaluated using distance-based classification techniques such as Manhattan distance, and it obtained 21.8% higher accuracy and enhanced privacy using keystroke-based authentication.

Shekhawat and Bhatt (2022) developed a user authentication system to analyze keystroke behavior using classification algorithms to model distinctive user profiles. However, a unique user behavior is modeled using SVM implementing inferential keystroke

archetype. It gathered the keystroke timely features from pressure sensors and are stored in an array. It aims to intensify the uniqueness of user profiles where diverse records for enhancing performance efficiency and evaluated using real-time dataset. The performance shows that SVM achieved a success rate of 97% for user authentication.

In contrast, Salem et al. (2023) recommended an online user authentication by a unique authentication method applying password and keystroke dynamics for enhanced comparative analysis. A unique prototype is proposed for gathering temporal and non-temporal information to extract the respective features. These features are essential in influencing the performance of authentication systems using keystroke dynamics by training Random Forest using these extracted features for estimating unique benchmark patterns. As a result, the performance of RF is superior to other existing solutions by obtaining EER and accuracy of 0% and 100%.

Furthermore, Raul et al. (2020) strengthened the keystroke dynamics by incorporating non-conventional features for analyzing static keystroke authentication. It enhances the performance in terms of FAR, FRR, EER and accuracy for analyzing non-conventional features along temporal information of users. It is evaluated using temporal information gathered from 30 users in terms of four diverse varieties of outcomes. i) SVM classifier analyzes only temporal features of keystrokes. ii) non-conventional features are analyzed through Logistic regression. iii) The SVM classifier is trained using a train-test split of 60% temporal and 40% non-conventional features. iv) The SVM classifier is trained using features containing 80% temporal information and 20% non-conventional information. Compared to other techniques, the performance of Logistic regression containing non-conventional features attained a lesser Equal Error Rate (EER) and higher accuracy of 90.5%. It is observed that utilization of temporal and non-conventional features shows better performance in authentication users through keystroke dynamics.

One of the major reasons for enhancing user authentication system is to combat spoofing attacks. Yunanto and Barmawi (2022) recognized the users by extracting features from the login process and compared against features from the keystroke database. The Indonesian anagram is utilized to evaluate by generating the words. It gathered word sets from 34 participants via smartphone keyboard. In addition, a combination of latency and

digraph plays a major role in bimodal keystroke dynamics that is useful for user authentication in smartphones by analyzing Euclidean distance. The performance of the proposed methodology witnesses that EER, FAR, and FRR are lowered by 23.075%, 16.381%, and 10.41% and enhanced the performance of the biometric authentication method for smartphone users.

However, Saini et al. (2020) examined a three-step authentication mechanism to validate a user with positions of sitting, walking, and relaxing while considering mobile orientations in both portrait and landscape. It integrates accelerometer data for classifying keystroke temporal information using Random Forest (RF) and K-Nearest Neighbour (KNN) classifiers. It obtained 2.9% EER for authenticating users while relaxing with mobile in landscape orientation. Subsequently, It is optimized through a heuristic approach such as Particle Swarm Optimization (PSO) to select optimal features for enriching the performance of RF and KNN using accelerometer data to examine three-step authentication. Optimized methods obtained a reduced EER of 2.2% for user authentication while in the position of relaxing and walking.

Some of the well-known feature selection techniques that are improved by researchers to optimize biometric authentication are defined below.

El-Kenawy et al. (2022) integrated an enhanced Dipper Throated Optimization technique using three search leaders to optimize the exploration technique, namely Dynamic Weighted Dipper Throated Optimization. It is validated using two datasets comprising keystroke features. Subsequently, a feature selection technique is proposed by choosing appropriate features for user authentication. The performance of the proposed DWDTO is compared with Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Gravitational Search Algorithm (GSA). However, Dipper Throated Optimization shows the improved performance for feature selection with BRNN in keystroke dynamics. BRNN achieved an accuracy of 99.32% and 99.02% for user authentication in both datasets.

On the other hand, Krishnamoorthy et al. (2018) employed Grid search optimization and wrapper method for parameter selection and feature selection to obtain the least

redundant and maximal relevant features *m*RMR with the aim of enhancing the classification and evaluated using performance metrics. It is concluded that significant parameters involving touch size, touch pressure, and coordinate positions are highly influential in detecting every user through analyzing key typing patterns of users with SVM-RBF. The performance of SVM incorporating the Radial Basis Function is observed to surpass the Linear function in recognizing user activities for keystroke authentication.

Alternately, Lamiche et al. (2019) devised a novel multi-modal for continuous authentication in a smartphone. It is accomplished by extracting gait and key typing patterns using an accelerometer from real-time data and continuous user intervention in both modalities. Subsequently, a fusion algorithm at the feature level is deployed to generate a multi-modal biometric user profile. The combined features are trained using a sequential floating forward technique to reduce feature dimensions. It is validated using a multi-modal dataset gathered in real-time through 20 subjects accomplishing diverse scenarios and trained using various machine learning techniques, including MLP, SVM, Random Forest, Random Tree, and Naïve Bayes. It is evident that MLP outperforms other techniques by obtaining 99.11% accuracy, 0.684% FAR, 1% EER and 7% FRR, respectively.

Additionally, several researchers used keystroke dynamics to develop user biometric authentication in Android mobile devices to intercept authentication attacks.

Thapliyal et al. (2022) proposed a novel user authentication approach by applying a hybrid technique and implemented in the Samsung On7 Pro C3590. The keystroke dynamics and their usefulness are explored for feature phones to offer an effective biometric framework. The security is enhanced in the feature phone by incorporating KNN and fuzzy logic to analyze the typing pattern of the user. The proposed mechanism is examined using keystroke features collected from 25 Samsung keyboard users and obtained an EER of 1.88%.

Meanwhile, Huang et al. (2020) implemented piezoelectric keystroke dynamics because piezoelectric effects detect the highly sensitive keystrokes and passive force, especially the most effective in incorporating cheap smartphones. The touch information estimated from user generated keystrokes comprised of touch and force frequencies for

specific users is retrieved from a piezoelectric touch panel as integral hardware for further processing using machine learning techniques such as SVM, Artificial Neural Network and RF. Random Forest attained an Equal Error rate of 0.72% which is comparatively lesser than other algorithms for user authentication and assures high security among smartphones.

In addition, Baynath et al. (2018) proposed and incorporated a novel machine learning algorithm, namely Neuro Evolution of the Augmenting Topology (P-NEAT). It analyze distinct keystroke temporal features and compare them with various machine learning techniques. It employes complex structure algorithms such as Fuzzy Expert System (FESs), NeuroEvolution of the augmenting topology (NEAT), Proposed NeuroEvolution of the augmenting topology, Support Vector Machine (SVM) and Chaotic Neural Network. The performance of these techniques is evaluated based on the Recognition Rate for analyzing the keystroke pattern of the user. It is evident that P-NEAT outperforms other techniques by achieving a 99% Recognition Rate and verifies users in terms of keystroke dynamics pattern.

Shi et al. (2021) exploited non-device techniques for user authentication by exploiting available WiFi signals obtained from IoT devices, including smart refrigerators, smart TVs, and smart thermostats. The WiFi signals are utilized to collect unique characteristics of the user based on human physiology and behavior in their day-to-day activities gathered while walking and standing. The features from channel state information are extracted from WiFi signals and are trained using deep-learning techniques for precise user authentication. Subsequently, a CNN-based deep learning technique is incorporated to handle signal distortion due to environmental causes and integrates features collected from multiple receivers. In addition, a transfer learning algorithm is employed to further analyze the daily activities in the proposed authentication system. The proposed system attained an accuracy of 94% for user authentication using 11 subjects based on WiFi activities.

Subsequently, Bhana and Flowerday (2020) suggested a two-tier user authentication mechanism using passphrases and keystroke dynamics with three theories, including Shannon entropy theory for examining password and passphrase strength, Chunking theory for assessing the issues of the passphrase and password memorization, and the Keystroke level model for assessing the issue of password and passphrase typing. It is evaluated using

two datasets for conducting login assessment to collect data in the form of passwords, passphrases, and expert review to verify and assess the research findings. It achieves a reduced false positive rate for enhancing security in user authentication.

In addition to Keystroke Static Authentication, Keystroke Continuous Authentication was widely utilized for user authentication mechanisms using edge computing (Chen et al., 2021). It achieved higher accuracy by implying anomaly detection using the Gaussian model and keystroke stream processing by providing enhanced keystroke profiles using static and continuous authentications. It is validated using three public datasets and case studies to authenticate the user in an online assessment. It attained FAR and EER of 0.05% and 3.43% for user verification using keystroke authentication to secure online examinations.

Alternately, Jadhav et al. (2017) exhibited the feature vectors such as mean and standard deviations for high-level security for users using keystroke dynamics and assessed in terms of FAR and FRR. Biometric operations such as user identification and verification are majorly focused on the proposed user authentication method. Unique pattern is gathered by estimating dwell time, and flight time using machine learning algorithms. Mean and standard deviation are applied for user authentication. It achieved FAR, FRR of 1%, and 4% with a 70% threshold value to differentiate valid and invalid users.

Meanwhile, Gu et al. (2021) implemented a novel authentication system using WiFi infrastructure to proposed WiPass Framework for analyzing keystroke dynamics gathered from channel state information to detect and authenticate spoofers. It incorporates a signal augmentation model known as Ricean distribution for intensifying user keystroke dynamics and applied a hybrid machine learning model comprising a CNN for extracting features and a SVM for classification. It utilizes a minimal, cost-effective single WiFi device for user authentication. The performance of hybrid model attained an accuracy of 92.1%, a false acceptance rate of 5.9%, and a false rejection rate of 6.3% for user authentication.

In contrast, the major focus of existing research is to enlarge the dataset by integrating three keystroke dynamic datasets collected from real-time users for user authentication available in existing studies.

In addition, Aversano et al. (2021) recommended the deep neural network to differentiate various users by analyzing the required feature set to generate unique key typing patterns. The optimization is accomplished by utilizing ensemble techniques containing super-classifiers and compared against diverse voting techniques involving majority and Bayesian along training strategies such as K-means and RF to choose the best classifier. K-means surpasses the RF classifier by obtaining an accuracy of 99.7% for achieving continuous authentication.

Pretrained models using machine learning and deep learning techniques play a major role in user authentication for keystroke dynamics by analyzing key typing features obtained from fixed-text passwords, namely .tie5Roanl.

Chang et al. (2022) trained the keystroke features using machine learning and deep learning techniques comprising Extreme Gradient Boosting (XGBoost), multi-layer perceptrons (MLP), K-NN, SVM, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and LSTM to find the unique keystroke pattern of a specific user for user authentication. It is observed that XGBoost attained a higher accuracy of 96.39% and outperforms other machine learning and deep learning techniques for biometric authentication using fixed-text keystroke typing characteristics.

On the other hand, Pentel (2017) aimed to analyze the traits of unintentional behavior gathered through external input devices such as keyboard and mouse for ultimately predicting the gender and age of users. The keystroke and mouse movement information is gathered from six various data sources available from 2011 to 2017, consisting of information from 1519 total subjects. This biometric data is analyzed using supervised machine learning techniques such as Logistic Regression, SVM, KNN, Decision Tree and RF. The performance of RF surpasses other techniques by obtaining a higher f-score to predict the gender and age of the anonymous users.

Additionally, Zeid et al. (2022) investigated the diverse probability of user classification by extracting features for keystroke dynamics from two diverse datasets. It comprises of various difficulty levels in fixed-text, without restriction on user typical behavior in free-text. Investigated and compared the performance of user authentication

using four classification algorithms involving RF, SVM, BayesNet (BN), and KNN to analyze fixed-text, and limitless continuous free-text dataset. The experimental result shows that Random Forest attained the highest accuracy for analyzing fixed-text and free-text at 98.8% and 87.58% and outperforms others in performing keystroke authentication.

In addition to supervised classification techniques, unsupervised algorithms are widely applicable for user authentication using keystroke dynamics.

Çevi□k et al. (2021) developed a web application-based user authentication approach that analyses the key typing behavior of users from 54 workers in an organization. It achieved successful user authentication by training the behavior model. The web application analyzes the collected keystroke data to estimate the unique keystroke archetype of the user by training machine learning techniques. The techniques include Tree based algorithms, decision tree and random forest. It is evaluated using benchmark data for user authentication. The performance of Tree based algorithms outperforms others by achieving the average of highest accuracy i.e. 94% for user keystroke authentication.

In contrast, Babu and Bhanu (2015) mentioned that machine learning techniques are effective for user authentication using keystroke dynamics.

Alternately, Chintalapudi et al. (2020) investigated the utilization of behavioral biometrics for user authentication by incorporating keystroke dynamics. The performance of keystroke authentication is enhanced by selecting the required feature using feature selection techniques for keystroke authentication. It applies classification techniques involving Multi-Layer Perceptron, K- Nearest Neighbor, and Support Vector Machine. It is evaluated using a keystroke benchmark dataset containing temporal information of keystrokes for a unique password ".tie5Roanl". The temporal data is gathered through 8 sessions where each user provides a password in 50 repetitions. In total, 400 keystroke data are collected, and the dataset contains 20400 records. The keystroke data is analysed using above mentioned three machine learning techniques to authenticate the user. As a result, the performance of MLP surpasses KNN and SVM by obtaining an accuracy of 90.93% for keystroke authentication.

On contrary, Singh et al. (2020) utilized the Boosting techniques for enhanced keystroke authentication using the CMU dataset. In particular, XGBoost is used to analyze the keystroke pattern of specified passwords using various classification techniques. The selected machine learning algorithm is incorporated in keystroke authentication system. It is evaluated with k-nearest neighbors, support vector classifier, and Random forest. The performance of XGBoost is higher by achieving 93.59% accuracy and outperforming other techniques for enhanced keystroke authentication.

Arsh et al. (2024) compared the performance of machine learning, deep learning and neural networks for user authentication based on keystroke dynamics. However, the performance of machine learning using RF obtained 99.9% accuracy and surpasses CNN, Deep Belief Network (DBN), Bidirectional-CNN (Bi-CNN) and Feed Forward Multilayer Neural Network (FFM-NN).

Meanwhile, Eltoukhy et al. (2024) integrated RF with whale optimization algorithm for its better feature selection and compared with Harris Hawks Optimization (HHO). It handles class imbalance problem using conditional tabular generative adversarial network (CTGAN) to generate synthetic data. It attained 99.62± 0.40% average accuracy and a reduction rate of 87.85% for effective smartphone authentication.

Sharma et al. (2025) experimented with multi-class CNN to obtain Keystroke Dynamic Images (KDI) for user authentication using both fixed-text and free-text keystroke dataset. These KDI features are analysed using SVM, DT and RF for multiclass classification. SVM obtained 50% accuracy, DT obtained 88%, and RF achieved 93% accuracy using KDI images. The combination of CNN and RF works better compared to CNN+SVM, and CNN+DT for user authentication.

Budžys et al. (2025) incorporated a framework that combines data fusion and deep learning techniques for enhanced keystroke authentication. The fixed-text keystroke datasets such as CMU, KeyRecs, GREYC-NISLAB datasets are integrated using three data fusion techniques namely linear, cubic and nearest neighbor. The fused data is converted into keystroke images using GAbor Filter MAtrix Transformation (GAFMAT). The images are analysed using Siamese neural network model for user authentication. The framework achieved 0.132% error rate for user authentication.

Putra and Chowanda (2025) enhanced the performance of continuous authentication using hybrid deep learning technique integrating CNN and Bi-directional Long Short Term Memory (BLSTM). In preprocessing, data cleaning, data transformation and feature extraction is performed to obtain keystroke features. These features are analysed using hybrid CNN-BLSTM technique. It obtained the minimal EER of 0.009%-0.127% for user authentication.

The following table 2.4 reviews biometric authentication using keystroke dynamics so far.

Table 2.4. Review of Recent Findings in Keystroke Dynamics

Study	Key findings	Algorithms applied	Observations
Putra and Chowanda (2025)	Integrated CNN and Bi-directional Long Short Term Memory for enhanced continuous authentication.	CNN-BLSTM	It obtained the minimum EER of 0.009% - 0.127%
Budžys et al. (2025)	Combined three data fusion techniques to aggregate the dataset, converted into images using GAFMAT, and analyzed using Siamese neural network for user authentication	Siamese neural network	It achieved lowest error rate of 0.132% for user authentication.
Sharma et al. (2025)	Combined CNN to extract Keystroke Dynamic Images and RF to analyze the KDI images for user authentication.	CNN, SVM, RF, DT	CNN+RF outperforms SVM and DT by attaining 93% accuracy using both fixed and free-text data.
Arsh et al. (2024)	Applied machine learning, deep learning and neural network for assessing user authentication performance.	RF, CNN, DBN, Bi-CNN and FFM-NN	RF outperforms and provides the best result in terms of 99.9% accuracy.

Study	Key findings	Algorithms applied	Observations
Eltoukhy et al. (2024)	Generated synthetic dataset using CTGAN for handling class imbalance problem and analyzed using random forest with WHO.	RF, whale optimization algorithm, Harris Hawks Optimization	RF with WHO obtained 99.62± 0.40% average accuracy higher than HHO for smartphone authentication.
Tewari & Verma (2022)	Applied CMU dataset and converted into image format in two CNN-pertained models, namely Resnet and AlexNet, and classified using SVM + PCA.	CNN, SVM + PCA	It is observed that ResNet outperforms and provides the best result in terms of 98% accuracy.
Sae-Bae & Memon, (2022)	The effectiveness of the individual verification using the keystroke dynamic is evaluated using proposed metrics in terms of FAR.	Manhattan distance, Euclidean distance, and Mahalanobis distance.	It is observed that less number of higher distinctiveness than lower distinctiveness and applied in all three classifiers based on the proposed distinctiveness score for keystroke dynamics is better to indicate the FAR.
Wang et al. (2022)	Proposed SIURUA, a novel authentication method utilizing keystroke and mouse dynamics along with calculating irrelevant features of scene and user for authentication.	Support vector machine	Obtained 84% accuracy and outperforms other state-of-art method.
Shekhawat & Bhatt (2022)	Developed a user authentication system using keystroke behavior to model unique profiles using classification algorithms.	SVM	Obtained 97% success rate using real-world dataset for authentication.
Salem et al.	Applied temporal and non-	Random Forest	The performance of RF

Study	Key findings	Algorithms applied	Observations
(2022)	temporal keystroke information for user authentication using an RF classifier.		surpassed other state of art methods in terms of EER and accuracy with 0% and 100%.
Yunanto & Barmawi (2022)	Challenged spoofing attacks by enhancing keystroke-based user recognition system using latency and digraph in the smartphone.	Euclidean distance	Obtained EER, FAR and FRR of 23.075%, 16.381%, and 10.41% and evaluated using smartphone user.
El-Kenawy et al. (2022)	Combination of Enhanced a Dipper Throated Optimization algorithm using three search leaders for feature selection for user authentication.	Dipper Throated Optimization	Achieved classification accuracy for two datasets utilizing optimized BRNN of 99.02% and 99.32% in terms of statistical analysis.
Thapliyal et al. (2022)	The proposed novel user authentication mechanism was implemented in the Samsung On7 Pro C3590 with 25 users.	KNN + fuzzy logic	It operates effectively and gained 1.88% EER.
Chang et al. (2022)	Machine Learning and Deep Learning for biometric authentication using the typing characteristics of fixed-text keystrokes.	XGBoost, multi-layer perceptrons (MLP), K-NN, SVM, CNN, RNN, LSTM	XGBoost outperforms and achieves the maximum accuracy at 96.39% for biometric authentication.
Zeid et al. (2022)	Investigated user authentication mechanism using four classification techniques based on difficult fixed-text and limitless continuous authentication.	Random Forest (RF), Support Vector Machines (SVM), BayesNet (BN), and K-Nearest	RF obtained the highest accuracy of 98.8% and 87.58% using fixed-text and free-text keystroke authentication and surpasses others.

Study	Key findings	Algorithms applied	Observations
		Neighbors (KNN)	
Chen et al. (2021)	Keystroke Static Authentication and Keystroke continuous authentication in online examination using three public data sets.	Gaussian model based anomaly detector and Sliding window,	It achieves 0.05% FAR and 3.43% ERR to verify users using keystroke dynamics.
Gu et al. (2021)	A protected keystroke authentication using a hybrid learning model with keystroke behavior including feature extraction and classification using CNN and SVM for WiFi authentication	CNN, SVM	Surpass maximum performance by obtaining accuracy, false acceptance and rejection rates of 92.1%, 5.9%, and 6.3%.
Aversano et al. (2021)	Proposed the ensemble-based DNN model to optimize the continuous authentication using keystroke dynamics and users allocated using best performing algorithm via Bayesian voting.	DNN, RF and k-means	K-Means outperformed the RF for user allocation and achieved an overall accuracy of 99.7% for Continuous authentication.
Shi et al. (2021)	The CNN based user authentication system using WIFI.	CNN	It achieves an accuracy of 94% using 11 subjects based on daily activities.
Maharjan et al. (2021)	The hybrid sensors (TENG+EMG) were proposed for keystroke authentication using a common password for user identification and authentication.	ANN, SVM	The performance of ANN is better than SVM and achieves 99% accuracy using keystroke verification.
Çevi□k et al. (2021)	Implemented web authentication based on user keystroke	Tree based algorithms,	Tree based algorithms obtained an average

Study	Key findings	Algorithms applied	Observations
	archetype applying machine learning algorithms and evaluated using 54 subjects.	decision tree, random forest	accuracy of 94% which is higher than others.
Andrean et al. (2020)	Explored deep learning algorithms based on Multi Layer Perceptron (MLP) for user authentication using CMU benchmark dataset.	MLP, scaled Manhattan, Mahalanobis Nearest Neighbor, Outlier Count, Neural Network	The performance of MLP suppresses others in terms of EER by obtaining 4.45% in user authentication.
Singh et al. (2020)	Analyzed the performance of various machine learning algorithms for keystroke-based user authentication using the CMU-CERT dataset.	XGBoost, k-nearest neighbors, support vector classifier, Random forest	XGBoost surpasses others based on accuracy by 93.59%.
Huang et al. (2020)	Proposed user authentication utilizing piezoelectric keystroke dynamics.	SVM, ANN and RF	RF outperforms other machine learning algorithms and achieves an EER of 72%.
Saini et al. (2020)	Examined a three-step authentication model considering mobile orientation and accelerometer data for optimization using Particle Swarm Optimization (PSO) in smartphones.	RF, K-Nearest Neighbour (KNN)	Optimized machine learning algorithms reduced EER by 2.2% from 2.9% for smartphone three-step authentication while comforting and walking positions.
Kim et al. (2020)	Proposed an innovative filter-based feature-selection method based on the trimmed mean and coefficient of variation for	Manhattan Distance	The proposed technique achieved 21.8% higher performance and surpasses other feature

Study	Key findings	Algorithms applied	Observations
	keystroke authentication utilizing a distance-based classification technique.		selection techniques.
Raul et al. (2020)	Suggested non-conventional features and timing data based on keystroke dynamics for user authentication.	Linear regression, random forest, gaussian, SVM	The performance of linear regression using non-conventional features obtained an accuracy of 90.5% and outperformed others.
Chintalapudi et al. (2020)	Investigated performance of MLP, KNN and SVM trained using keystroke archetype for user authentication.	Multi Layer Perceptron(MLP, KNearest Neighbor (KNN), Support Vector Machine (SVM)	The performance of MLP outperforms KNN and SVM in terms of accuracy with 90.93%, 37%, and 76% using the CMU CERT dataset.
Bhana & Flowerday (2020)	Suggested a two-tier user authentication protocol that used keystroke and paraphrasing to improve security.	SET and CT	Reduced the FPR.
Baynath et al. (2018)	Explored various machine learning algorithms for recognizing the keystroke pattern of users.	FES, NEAT, P-NEAT, SVM and Chaotic Neural Network	P-NEAT outperforms other algorithms and achieves a 99% Recognition Rate (RR).
Pentel (2017)	Classification of mouse and keystroke data from six different data sources to predict anonymous user age and gender	Logistic Regression, SVM, KNN, C4.5, RF	RF performs better than other algorithms, but the result is still preliminary.
Jadhav et al. (2017)	The proposed method exhibits a high level of security for user authentication using keystroke dynamics.	Mean and Standard Deviation	It accomplishes 1% FAR and 4% FRR, respectively.

From the literature, it is observed only literature focuses on general keystroke dynamics is available because no authors employed keystroke dynamics for unintentional insider mitigation. Most of above studies applied metaheuristic approach for optimizing keystroke authentication. A feature engineering technique is required for keystroke dynamics.

2.4. RESEARCH GAPS IDENTIFIED

The literature study shows the following research gaps identified and the analysis of research gap identified is shown in table 2.5.

- Boosting algorithms fall short for insider threat detection due to data irregularity.
- Class imbalance problem is pervasive and not majorly focused.
- Very less focused on detecting both intentional and unintentional insiders with minimal misclassification rates. So, an attempt has been made to detect both intentional and unintentional insider threats.
- The researchers show minimal interest in unintentional insider mitigation. Only specified the effective mitigation strategies for unintentional insiders.
- Not much importance is given to feature engineering techniques for optimizing deep learning techniques which effectively mitigate intentional insiders.
- In existing research, mitigation of intentional insider is not focused. Hence, an attempt is made to mitigate the intentional insiders.

Table 2.5: Analysis of Research Gap Identified and Proposed Contribution

Area	Research Gaps	Proposed contribution
Intentional Insider Detection	Class Imbalance Problem	Enhanced Nearmiss2 Sampling
	Increased Misclassification Rate	Hybrid machine learning solution
Unintentional Insider Detection	No solution for detecting both intentional and unintentional insider threat	Hybrid machine learning solution

Unintentional insider mitigation	Only explored feature engineering techniques using metaheuristic approach for optimizing user authentication	Proposed enhanced feature engineering technique for behavioral biometrics
Intentional insider mitigation	Mitigation was not focused	User profiling mechanism

2.5. CHAPTER SUMMARY

From the literature, it is obvious that there is no single comprehensive methodology for detecting and mitigating both intentional and unintentional insiders at present. Existing research focus less on tuning the sampling techniques to better handle the class imbalance problem. Detection of both intentional and unintentional insiders with more accuracy and minimal misclassification rate is required. So far now, very few focus on mitigation of both intentional and unintentional insider threats. Hence, an attempt has been made to develop a comprehensive methodology for detecting and mitigating both intentional and unintentional insiders. The next chapter discusses the proposed methodology.