

Email-Based Internal Threat Detection using Machine Learning

Neha Santosh Jadhav¹ and Ms. Yogita Kadbane²

¹Student, Master of Computer Application

²Assistant Professor, Master of Computer Application

Sadhu Vaswani Institute of Management Studies for Girls, Pune

Savitribai Phule Pune University, Pune, Maharashtra, India

Abstract: Employees who have authorized access to an organization's systems and data are allowed to perform Employee work activities. However when an Employee performs activities or violates security policies the behaviour may indicate a potential insider threat. Employee email can provide information for identifying such behaviour because it contains details such as the sender, recipient, timestamp, attachments, message size, activity type and action. In this research the SPEDIA email dataset is used to analyse Employee email activities and identify whether the activities are normal or anomalous using Machine Learning techniques.

The study uses 14,562 training records and 3,641 testing records with 24 attributes in the SPEDIA dataset. During preprocessing the required categorical and text-related information is converted into form resulting in 26 features, for Machine Learning. Three algorithms, Logistic Regression, Decision Tree and Random Forest are trained and compared using Accuracy, Precision, Recall and F1-Score. Among the three models Decision Tree gives the overall performance with 91.02% Accuracy, 92.01% Precision, 95.95% Recall and 93.94% F1-Score.

The proposed system provides a way to monitor Employee email activities and identify activities that appear unusual. Of manually checking a large number of email records Machine Learning can analyse the available information and help identify potentially risky activities. The proposed system can save time.

Keywords: Insider Threat using Machine Learning, Email Security, Anomaly Detection, Decision Tree, Cyber Security, SPEDIA Dataset, Employee Behaviour Threat Detection, Cyber Threat Detection.

I. INTRODUCTION

In today's world every company uses computers, networks and email for their everyday tasks. Many workers are allowed to access the company systems and data. Because of this sometimes workers might do something dangerous using their company access. This kind of issue is known as a threat. It is hard to find a threat because strange actions from workers can sometimes look like regular work. Email can be useful in finding these kinds of worker actions because it has details like: To, From, Date and time Attachments, Email size, Activity, Action. In this project the SPEDIA email dataset is used to figure out if an employee's email activity is normal or strange. To find patterns in the data I used machine learning models to predict email activities.

This project includes three Machine-Learning algorithms: Logistic Regression, Decision Tree and Random Forest. These were used to compare and pick the model for the final detection. For insider threats it can be hard to find them because workers usually have access to company resources. A strange email might not always mean something but it can be an early sign of a problem. Looking at email activities, like sending patterns, attachments, people who get the emails, message size and other details can help spot dangerous behavior. Machine Learning can simplify this process by learning patterns from previous email activities and using them to detect unusual activities in new data.



The main aim of this research is to develop a simple machine-learning-based approach for detecting anomalous employee email activity. Using the SPEDIA dataset, this research trains and tests Logistic Regression, Decision Tree, and Random Forest models and compares their performance using Accuracy, Precision, Recall, and F1-Score. The final system classifies email activities as Normal or Anomalous and provides useful information for further security investigation. By focusing on early detection of potentially risky behaviour, this approach offers a practical way for organization to monitor employee email activity.

II. LITERATURE REVIEW

Many researchers have looked into using Machine Learning for insider threat detection. For looking at employee behaviour. Insider threats happen when an employee who has permission to access systems does something that's not usual or that is risky. Different studies have taken a look at employee behaviour, system logs, email communication and other records of activity to find actions. The SPEDIA dataset was created for research on insider threats. Holds different kinds of employee activity data including information about emails.

Email data is helpful for this kind of research because it has details like who sent the message, who received it the time it was sent any attachments, the size of the message, the type of activity and what action was taken. So datasets like SPEDIA offer useful information for studying Machine Learning-based insider threat detection. Different Machine Learning algorithms have been used to find employee activities. Logistic Regression is often used because it is simple and easy to set up.

Decision Tree is helpful because the choices it makes are easier to understand while Random Forest uses decision trees to make the prediction process better. Other studies have also used advanced Machine Learning and Deep Learning methods along with ways to select features and find anomalies. These studies show that Machine Learning can help find employee activities that're not normal. Based on this research the current study uses the SPEDIA email dataset and compares Logistic Regression, Decision Tree and Random Forest to classify employee email activities, as Normal or Anomalous.

Research Gap

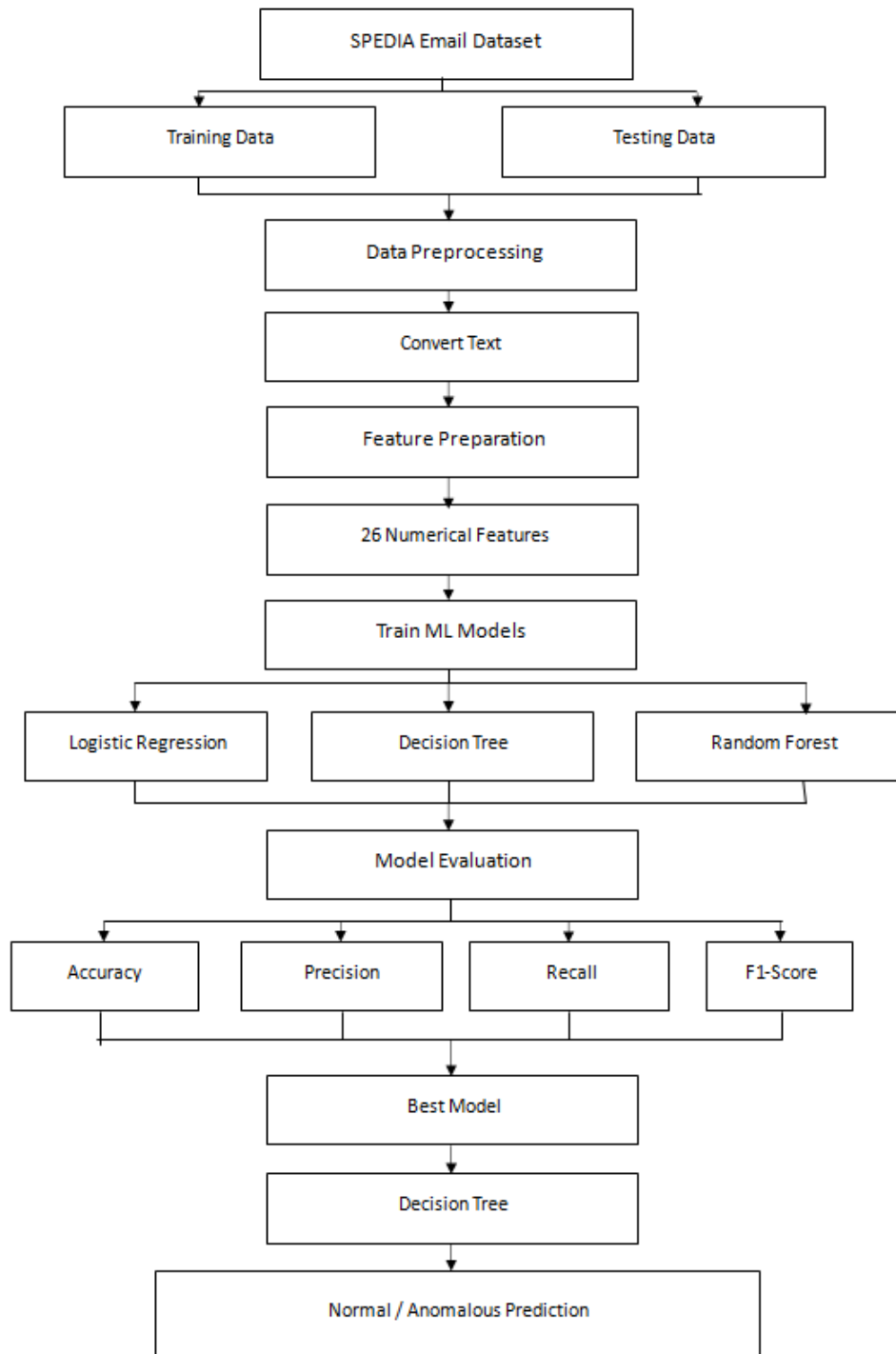
Existing studies have used machine learning to detect the insider threats, which mainly focuses on general employee activities rather than email-specific activities. There is a need to analyze important email details like From, To, Time, Attachments, Size, Activity, Action from this it is easy to identify unusual behaviour. Also if checking a large amount of email data manually it becomes difficult and also time-consuming.

III. METHODOLOGY

The methodology consists of following main steps:

Dataset collection, Data Preprocessing, Feature encoding, Train ML models, Model evaluation, Best model.





Dataset

The SPEDIA dataset is used in this project, it contains both normal and anomalous data .

The project uses: 14,562 training records,
3 641 testing records.

Data Preprocessing

The SPEDIA email dataset that contains different types of data. In that some columns contain numbers while other columns contain text data such as user name, email, addresses, activities, actions and so on.

Machine Learning models cannot directly understand this type of text information. Therefore, before training the models , the data must need to be cleaned and converted into a numerical format.

In this step, data was cleaned and prepared for machine learning models. The text information is converted in numerical data , so that models could understand them. This data holds important details like user, email, time, attachments, size, activity, and action, these all were used to make a prediction.

The original dataset contained 24 columns. After preprocessing the data it creating the required numerical features, after this process the data was converted into 26 model features.

IV. MACHINE LEARNING MODELS

Three different machine learning algorithms are selected for this research:

Random Forest – Random Forest is a machine learning algorithm that uses decision trees to make final prediction. Of relying on one tree it combines the results of multiple trees. It can handle a number of network features and also complex data. This makes the model reduce prediction.

Decision Tree – Decision tree is an algorithm that makes decision step, by step. It works like a flowchart. It can classify traffic normal or different types of attack.

Logistic Regression – Logistic Regression is a machine learning algorithm used for classification. It checks the input features. Calculates probability. It is easy to use and requires computing power.

Algorithm

1. Start
2. load the SPEDIA email dataset
3. Check and clean the data
4. Select Useful email and activity features
5. Convert text data into number
6. Use the testing data to make prediction
7. Check Accuracy, Precision, Recall, F1-Score
8. Compare all three models and select the best model - Decision Tree
9. Use Decision Tree to Classify email activity as Normal or Anomalous.

V. MODEL EVALUATION

The models are evaluated using:

Accuracy

Accuracy is use to show how correctly the model identifies the email activity is Normal or Anomalous from the testing data.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision

Precision that shows how many email are predicted as Anomalous are actually a anomalous or not. It helps to reduce false alerts.



$$\text{Precision} = \text{TP} / \text{TP} + \text{FP}$$

Recall

Recall that shows how many actual anomalous email activities are correctly detected by the model. Fewer threats are missed means higher recall.

$$\text{Recall} = \text{TP} / \text{TP} + \text{FN}$$

F1-Score

F1-Score shows that the balance between Precision and Recall can helps to select the best performing model.

$$\text{F1} = 2 \times \text{Precision} \times \text{Recall} / \text{Precision} + \text{Recall}$$

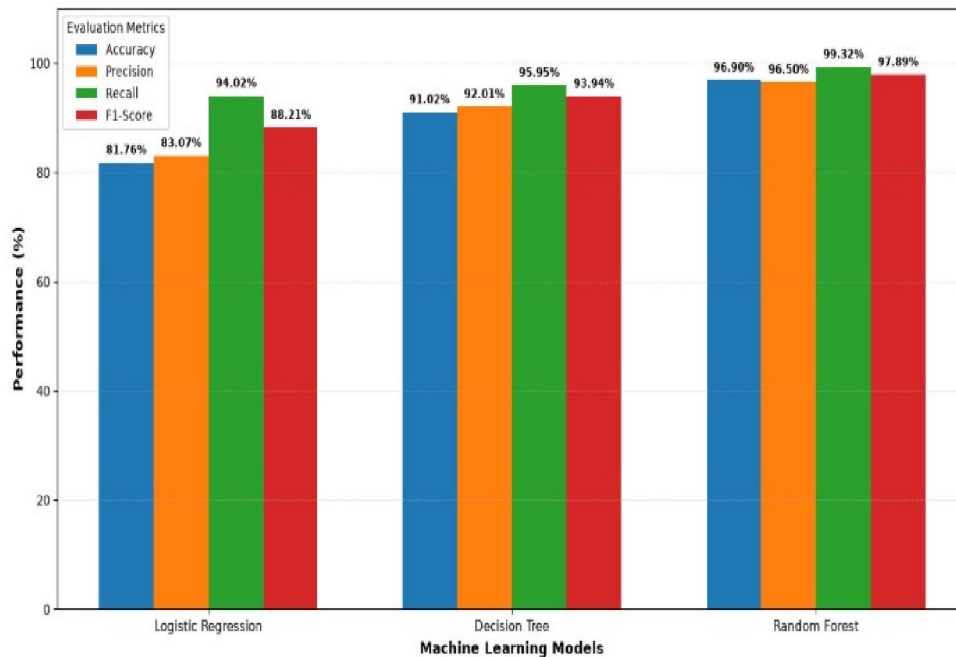
Finding and Results

Model Performance

After training and testing the Three models, the results were:

	Model	Accuracy	Precision	Recall	F1-Score
0	Logistic Regression	66.99	71.26	91.33	80.05
1	Decision Tree	91.02	92.01	95.95	93.94
2	Random Forest	87.70	92.32	90.57	91.44

Machine Learning Model Performance Comparison



VI. MAJOR FINDINGS

This study shows that the machine learning can help to detect unusual employee email activities. In this, the SPEDIA dataset contains 14,562 training records and 3,641 testing records. The original dataset has 24 columns which holds the details like User, From, To, Time, Attachments, Size, Activity, and Action. Before using the machine learning the data was cleaned and prepared for it. After preprocessing the model 26 numerical features were prepared. Three models were tested: Logistic Regression, Decision Tree, and Random Forest.

These three models were compared using Accuracy, Precision, Recall, and F1-Score. So from this Decision Tree was selected as a best model. From this the model can classify employee email activities as a Normal or Anomalous. This helps to identify unusual email activities more quickly.

VII. LIMITATION

This research has some limitations that should be considered while using this model. This study only uses the SPEDIA email dataset, so the result may be different when it is applied to other organization or datasets. This mainly focuses only on employee email activities and it does not cover all types of insider activities. The model can classify activities as a Normal or Anomalous, but it cannot confirm whether an employee is actually malicious.

VIII. CONCLUSION

This research developed a successful machine-learning based system for detecting unusual employee email activities by using SPEDIA email dataset. In this, the system studies how different email and activity details can help to classify them as a normal or anomalous. Therefore it makes easier to identify any activity that may need further attention.

In this we used to study 14,562 training records and 3,641 testing records. In the original dataset it has 24 attributes, which were now processed to prepare a 26 numerical features. In this, model we are using three machine learning algorithms: Logistic Regression, Decision Tree, Random Forest- these algorithms were trained and to compare Accuracy, Precision, Recall, F1-Score.

Among them the tested models, Decision Tree performed the best in the selected experiment, with 91.02% Accuracy, 92.01% Precision, 95.95% Recall, 93.94% F1Score. So in, these the result shows that the model can identify effectively a large number of potential anomalous activity.

Overall, this system helps to find unusual employee email activity early. It shows details like To, From, Time Attachment, Activity, and Action. From this data it helps the security team to check suspicious email. However, an anomalous result does not always mean the all employee is a threat, it only means whatever the activity looks like a unusual and it need to be check

REFERENCES

1. Malvika Singh, B. M. Mehtre & S. Sangeetha(2020)- Insider Threat Detection Based on User Behaviour Analysis.
2. S. Asha, D. Shanmugapriya & G.Padmavathi(2023)- Malicious Insider Threat Detection Using Variation of Sampling Method For Anomaly Detection .
3. Pennada, S. S. P., Nayak, S. K., & Vamsi Krishna, M.(2025)- Insider Threat Detection Using Behavioural Analysis through Machine Learning and Deep Learning Techniques
4. Mohd Anul Haq, Mohd Abdul Rahim Khan & Mohammed Alshehri(2022)- Insider Threat Detection Based on NLP Word Embedding and Machine Learning.
5. Colin Chibaya & Ibidun Christiana Obagbuwa(2025)- Machine Learning and Sentiment Analysis For Predicting Insider Threats on Email.
6. Antonia Michael (2020) – A Machine Learning Approach to Detect Insider Threats in Emails Caused by Human Behaviour.



7. Naan Kang Garba, Sandip Rakshit, Dakun Mang Chai & Narasimha Rao Vajjhala (2021) – An Email Content-Based Insider Threat Detection Model Using Anomaly Detection Algorithms.
8. .David Álvarez Muñoz, Luis Perez Miguel, Alberto Mateo Muñoz, Xavier Larriva-Novo, Manuel Alvarez-Campana & Diego Rivera (2026) – Design and Generation of a Dataset for Training Insider Threat Prevention and Detection Models: The SPEDIA Dataset.

