***Algorithm for insider threat detection.***

Step1: Data gathering: - Gathering email dataset of an organization. Both the data email sent & email received.

Step2: Preprocessing: - Cleaning, tokenizing and vectorizing the text data. This step is crucial and important too because this is the data which will serve as an input for machine learning model and a deciding factor to identify the type of threat.

Step3: Feature extraction: - Sender and recipients address, timestamps, email content, attachment, metadata.

Step4: Labelling data: - To train the machine learning model we need some emails which consist of threat and which are threat free.

Step5: Model selection: - E.g. 1. Logistic regression

2. Random Forest

3. Support Vector Machine

4. Gradient boosting machine

5. Deep learning model (LSTM, CNN) for text classification

**Or any other.**

Step6: Training the Model: - Train the machine learning model using the labelled dataset.

Step7: Evaluation: - Evaluate the performance of the trained model using terminologies such accuracy, precision, F1 score, and many more as per requirement. Ensure that the model performs well on training data and test data.

Step8: Deployment: -Deploy the trained the ml model into your email infrastructure/dataset for real time insider threat detection.

**Process to doo**

1. Feature extraction us