**BSCIT-05-0104/2022**

**FLORIDA CHEPKURUI KORIR**

**ASSIGNMENT 1: DATA SCIENCE**

**1. Approach to Analyzing Insider Threats in a Financial Institution**

To detect potential insider threats, I would take a systematic, multi-layered approach to analyzing the diverse data sources while balancing security and privacy concerns.

**Step 1: Data Collection & Preprocessing**

* **Network Logs:** Analyze unusual data transfers, login attempts, and access to restricted systems.
* **Employee Access Records:** Detect anomalies in system access outside of normal working hours or unauthorized attempts.
* **Email Communications:** Apply Natural Language Processing (NLP) to flag sensitive keywords related to fraud, data exfiltration, or internal complaints.

**Step 2: Feature Engineering & Anomaly Detection**

* **Behavioral Profiling:** Establish normal baselines for each employee’s activities.
* **Unsupervised Learning:** Use clustering (e.g., K-Means, DBSCAN) to detect deviations.
* **Graph Analysis:** Identify hidden relationships between employees and sensitive data.
* **Time-Series Analysis:** Detect abnormal spikes in access patterns.

**Step 3: Addressing Privacy & Ethical Considerations**

* **Data Minimization:** Analyze only metadata rather than full content where possible.
* **Access Control:** Restrict investigation data to authorized cybersecurity personnel.
* **Transparency:** Inform employees about monitoring policies to maintain trust.

**Step 4: Communicating Findings**

* **Technical Teams:** Provide dashboards with detailed logs and anomaly scores.
* **Non-Technical Stakeholders:** Present risk-based summaries with visualizations.