

PS 12 : Situated Learning Document Submission Channl.

Name : RAJASEKARAN . S

BITS ID : 2024AD05043

GROUP # : 305

Company : Citi Bank

Department : Third Party Sourcing Management & Payments

Role : Business Analyst

Q1: Vendor Spend Anomaly Detection

Problem Statement:

Am part of a team managing ~~large~~ <sup>huge</sup> number of vendors across multiple departments (procurement, services, operators etc.). Volume of payment will be humongous, thereby the risk of duplicate and incorrect payments were also high. Manual audit was not feasible due to the volume. Hence detecting the anomaly were often happening either delayed or very slow in performance.

Goal : To monitor the payments Real-Time & identify any,

- \* Suspicious Payments
- \* Duplicate Payments

Challenges :

- \* Multiple data sources, with different data format
- \* No common pattern across all the vendors. Few are high value vendors, so setting a common threshold is not feasible.

How it was addressed ?

- \* Identified the key fields from the source data.
- \* Feature extraction on the payments per certain frequency (day, month, week).
- \* Historical thresholds by vendor per category / region.
- \* Detected the anomaly using one local outlier factor model.
- \* Based on the transaction behavior and its respective clustering, identified the local outliers.

Q2 : Limitations of the Implemented Model:

- \* for minor variations, more of false Positives.
- \* Not fully Real Time. kind of Batch Processing.
- \* Not good in scaling much high volume of transactions.

How to Modify :

We are analysing on below options, sourced from Internet and do the enhancement of the existing model and use for the new developments as well.

- Scaling to Approximate Nearest Neighbors.
- Real-time inference with Pre-trained models.
- Switching to Streaming-compatible Models.