

## INTISARI

### **PENGEMBANGAN ALGORITMA PRA PEMROSESAN UNTUK MENDUKUNG DETEKSI ANOMALI BERBASIS K-MEANS PADA AKTIVITAS TRANSAKSI *PAYMENT AGGREGATOR***

Oleh

Reka Alamsyah

22/509199/PPA/06452

*Sistem payment aggregator dengan fitur single settlement meningkatkan efisiensi transaksi. Namun demikian hal ini berisiko terhadap serangan siber dan kesalahan system. Resiko ini akan memunculkan kejadian abnormal atau anomali. Middleware service mencatat aktivitas transaksi dalam bentuk log. Data log dapat dianalisis untuk deteksi anomaly sebagai akibat dari serangan siber ataupun kesalahan system. K-Means clustering kurang efektif untuk mendeteksi anomaly pada data log karena data log transaksi sering kali tidak terstruktur, tidak konsisten, dan memiliki skala fitur yang bervariasi. Penelitian ini mengembangkan algoritma pra pemrosesan untuk meningkatkan kualitas data sebelum klasterisasi. Data log transaksi dari Juli hingga Desember 2023 digunakan, dengan tahapan Pra Pemrosesan mencakup normalisasi, standarisasi, dan Principal Component Analysis (PCA). K-Means diterapkan dengan inisialisasi K-Means++, dan jumlah klaster ditentukan menggunakan kneedle algorithm. Hasil menunjukkan bahwa standarisasi tanpa PCA meningkatkan segmentasi yang memiliki nilai DBI rata-rata 0.2 dan efektifitas deteksi anomali dengan nilai recall yang berikan memiliki rata-rata 98%.*

**Kata Kunci:** *Preprocessing, K-Means, Deteksi Anomali, Log Middleware, Kneedle Algorithm*

## ABSTRACT

### ***ALGORITHM DEVELOPMENT FOR PREPROCESSING TO SUPPORT ANOMALY DETECTION BASED ON K-MEANS IN PAYMENT AGGREGATOR TRANSACTION ACTIVITIES***

by

Reka Alamsyah

22/509199/PPA/06452

*The payment aggregator system with a single settlement feature improves transaction efficiency. However, it also introduces risks of cyberattacks and system failures, which may lead to abnormal events or anomalies. Middleware services record transaction activities in the form of logs. These logs can be analyzed to detect anomalies resulting from cyberattacks or system errors. K-Means clustering is often ineffective for anomaly detection in log data due to its unstructured format, inconsistent structure, and varying feature scales. This study develops a preprocessing algorithm to improve data quality prior to clustering. Transaction log data from July to December 2023 was used, with preprocessing steps including normalization, standardization, and Principal Component Analysis (PCA). K-Means clustering was applied using K-Means++ initialization, and the number of clusters was determined using the kneedle algorithm. The results show that standardization without PCA enhances segmentation, achieving an average Davies-Bouldin Index (DBI) of 0.2 and anomaly detection effectiveness with an average recall of 98%.*

**Keywords:** *Data Preprocessing, K-Means, Anomaly Detection, Middleware Logs, Kneedle Algorithm*