

INTISARI

DETEKSI ANOMALI PADA KLAIM BIAYA RUMAH SAKIT KEPADA BPJS KESEHATAN MENGGUNAKAN METODE DENSITY CANOPY K-MEANS

Hendri Kurniawan Prakosa
19/448706/PPA/05789

BPJS Kesehatan yang telah berdiri selama hampir satu dekade masih mengalami defisit dalam proses penjaminannya terhadap peserta. Salah satu faktor yang menyebabkan hal tersebut adalah adanya ketidaksesuaian pada proses klaim yang cenderung merugikan. Misalnya saja dengan menaikkan *coding* diagnosa sehingga klaimnya menjadi lebih besar, melakukan klaim ganda atau bahkan pencatatan klaim palsu. Tindakan-tindakan tersebut berdasarkan peraturan pemerintah termasuk pada tindakan kecurangan (*fraud*). Tindakan *fraud* dapat dideteksi dengan melihat adanya anomali yang muncul pada data klaim.

Penelitian ini bertujuan untuk mengetahui anomali data klaim rumah sakit. Data yang digunakan adalah data klaim BPJS tahun 2015-2016. Untuk mengetahui anomali data tersebut digunakan Algoritma K-Means. Agar klaster yang terbentuk optimal digunakan metode density canopy pada penentuan centroid awalnya. Pada proses preprocessing data, digunakan Principal Component Analysis (PCA) untuk mereduksi variabel data dan normalisasi dengan menggunakan z-score. Berdasarkan score variasi dari PCA variabel biaya tagih dan lama dirawat paling berpengaruh pada dataset.

Evaluasi pembentukan klaster menggunakan silhouette index menghasilkan capaian nilai sebesar 0.82 dengan jumlah klaster 5 dan nilai RMSE hasil pemodelan dengan *simple linear regression* sebesar 0.49 untuk biaya tagih dan 0.97 untuk variabel lama dirawat. Berdasarkan hasil pemodelan tersebut terlihat titik anomali sebanyak 435 dari 10.000 data atau sebesar 4.35 % dan klaster ke-5 terkonfirmasi sebagai klaster anomali. Diharapkan dengan diketahuinya anomali tersebut dapat dilakukan tindak lanjut yang lebih efektif.

Kata Kunci: Deteksi anomali, BPJS Kesehatan, *density based*, K-Means

ABSTRACT

ANOMALY DETECTION IN HOSPITAL COST CLAIMS TO BPJS KESEHATAN USING THE DENSITY CANOPY K-MEANS METHOD

BPJS Kesehatan, which has been in existence for almost a decade, is still experiencing a deficit in the process of guaranteeing participants. One of the factors that causes this is a discrepancy in the claims process which tends to be detrimental. For example, by increasing the diagnostic coding so that the claim becomes bigger, making multiple claims or even recording false claims. Based on government regulations these actions are including fraud. Fraud can be detected by looking at the anomalies that appear in the claim data.

This study aims to determine the anomaly of hospital claims data. The data used is BPJS claim data for 2015-2016. To find out the data anomalies, the K-Means Algorithm is used. In order for the optimal cluster to be formed, the density canopy method was used to determine the initial centroid. In the data preprocessing process, Principal Component Analysis (PCA) is used to reduce data variables and normalize using z-score. Based on the variation score of the PCA, the variable billing fee and length of stay has the most effect on the dataset.

Evaluation of cluster formation using a silhouette index with an achievement value of 0.82 with the number of clusters 5 and the RMSE value of modeling results with simple linear regression of 0.49 for billing fee and 0.97 for length of stay. Based on that, there are 435 anomalous points out of 10.000 data or 4.35% and the 5th cluster is confirmed as an anomaly cluster. It is hoped that by knowing these anomalies, more effective follow-up can be carried out

Keywords: Anomaly detection, BPJS Kesehatan, density based, K-Means